THE LIMITS OF PRICE DISCRIMINATION

By

Dirk Bergemann, Benjamin Brooks, and Stephen Morris

May 2013
Revised September 2014

COWLES FOUNDATION DISCUSSION PAPER NO. 1896RRR

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

The Limits of Price Discrimination

Dirk Bergemann, Benjamin Brooks, and Stephen Morris*

September 24, 2014

Abstract

We analyze the welfare consequences of a monopolist having additional information about consumers' tastes, beyond the prior distribution; the additional information can be used to charge different prices to different segments of the market, i.e., carry out "third degree price discrimination". We show that the segmentation and pricing induced by the additional information can achieve *every* combination of consumer and producer surplus such that: (i) consumer surplus is non-negative, (ii) producer surplus is at least as high as profits under the uniform monopoly price, and (iii) total surplus does not exceed the surplus generated by efficient trade.

1 Introduction

A classic issue in the economic analysis of monopoly is the impact of discriminatory pricing on consumer and producer surplus. A monopolist engages in *third degree price discrimination* if he uses additional information about consumer characteristics to offer different prices to different segments

*Bergemann: Department of Economics, Yale University, Box 208281, New Haven, CT 06520-8281 (email: dirk.bergemann@yale.edu); Brooks: Becker Friedman Institute, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637 (email: babrooks@uchicago.edu); Morris: Department of Economics, Princeton University, Fisher Hall, Princeton, NJ 08540 (email: smorris@princeton.edu). We thank the co-editor, Andy Skrzypacz and three anonymous referees for many helpful and productive suggestions. We would like to thank Nemanja Antic, Simon Cowan, Emir Kamenica, Babu Nahata, Omer Reingold, Lars Stole, Juha Tolvanen, and Ricky Vohra, as well as many seminar participants, for informative discussions. Lastly, thanks to Alex Smolin and Áron Tóbiás for excellent research assistance. We gratefully acknowledge financial support from NSF SES 0851200 and ICES 1215808. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

of the aggregate market. A large and important literature (reviewed below) examines the impact of particular segmentations on consumer and producer surplus, as well as on output and prices.

In this paper, we characterize what could happen to consumer and producer surplus for all possible segmentations of the market. We know that at least two points will be attained. If the monopolist has no information beyond the prior distribution of valuations, there will be no segmentation. The producer charges the uniform monopoly price and gets the associated monopoly profit, which is always a lower bound on producer surplus; consumers receive a positive surplus, the standard information rent. This is marked by point A in Figure 1. On the other hand, if the monopolist has complete information about the valuations of the buyers, then he can charge each buyer their true valuation, i.e., engage in perfect or *first degree price discrimination*. The resulting allocation is efficient, but consumer surplus is zero and the producer captures all of the gains from efficient trade. This is marked by point B in Figure 1.

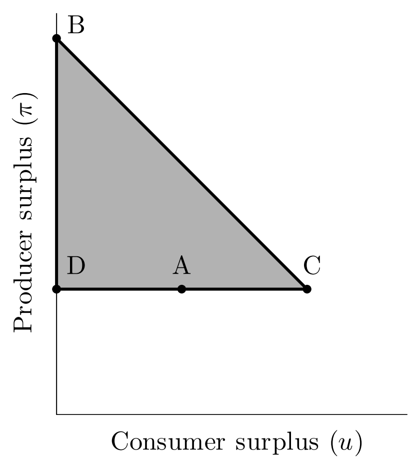

Figure 1: The surplus triangle

We are concerned with the welfare consequences of all possible segmentations, in addition to the two mentioned above. To begin with, we can identify some elementary bounds on consumer and producer surplus in any market segmentation. First, consumer surplus must be non-negative as a consequence of the participation constraint; a consumer will not buy the good at a price above his valuation. Second, the producer must get at least the surplus that he could get if there was no segmentation and he charged the uniform monopoly price. Third, the sum of consumer and producer surplus cannot exceed the total value that consumers receive from the good, when that value exceeds the marginal cost of production. The shaded right angled triangle in Figure 1 illustrates these three bounds.

Our main result is that *every* welfare outcome satisfying these constraints is attainable by some market segmentation. This is the entire shaded triangle in Figure 1. The point marked C is where consumer surplus is maximized; in particular, the producer is held down to his uniform monopoly profits, but at the same time the outcome is efficient and consumers receive all of the gains in efficiency relative to no discrimination. At the point marked D, social surplus is minimized by holding producer surplus down to uniform monopoly profits and holding consumer surplus down to zero.

We can explain these results most easily in the case where there is a finite set of possible consumer valuations and the cost of production is zero. The latter is merely a normalization of constant marginal cost we will maintain in the paper. We will first explain one intuitive way to maximize consumer surplus, i.e., realize point C. The set of market prices will consist of every valuation less than or equal to the uniform monopoly price. Suppose that we can divide the market into segments corresponding to each of these prices in such a way that (i) in each segment, the consumers' valuations are always greater than or equal to the price for that segment; and (ii) in each segment, the producer is indifferent between charging the price for that segment and charging the uniform monopoly price. Then the producer is indifferent to charging the uniform monopoly price on *all* segments, so producer surplus must equal uniform monopoly profit. The allocation is also efficient, so consumers must obtain the rest of the efficient surplus. Thus, (i) and (ii) are sufficient conditions for a segmentation to maximize consumer surplus.

We now describe a way of constructing such a market segmentation iteratively. Start with a "lowest price segment" where a price equal to the lowest valuation will be charged. All consumers with the lowest valuation go into this segment. For each higher valuation, a share of consumers with that valuation also enters into the lowest price segment. While the relative share of each higher valuation (with respect to each other) is the same as in the prior distribution, the proportion of all of the higher valuations is lower than in the prior distribution. We can choose that proportion between zero and one such that the producer is indifferent between charging the segment price and the uniform monopoly price. We know this must be possible because if the proportion were equal to one, the uniform monopoly price would be profit maximizing for the producer (by definition); if the proportion were equal to zero—so only lowest valuation consumers were in the market—the lowest price would be profit maximizing; and, by keeping the relative proportions above the lowest valuation constant, there is no price other than these two that could be optimal. Now we have created one market segment satisfying properties (i) and (ii) above. But notice that the consumers not put in the lowest price segment are in the same relative proportions as they were in the original population, so the original uniform monopoly price will be optimal on this "residual segment". We can apply the same procedure to construct a segment in which the market price is the second lowest valuation: put all the remaining consumers with the second lowest valuation into this market; for higher valuations, put a fixed proportion of remaining consumers into that segment; choose the proportion so that the producer is indifferent between charging the second highest valuation and the uniform monopoly price. This construction iterates until the segment price reaches the uniform monopoly price, at which point we have recovered the entire population and point C is attained.

For the formal proof of our results, we make use of a deeper geometric argument. This establishes an even stronger conclusion: any point where the monopolist is held down to his uniform monopoly profits—including outcomes A, C, and D in Figure 1—can be achieved *with the same segmentation*! In this segmentation, consumer surplus varies because the monopolist is indifferent between charging different prices. A geometric property is key to the existence of such a segmentation. Consider the set of all markets where a given monopoly price is optimal. This set is convex, so any aggregate market with the given monopoly price can be decomposed as a weighted sum of markets which are extreme points of this set, which in turn defines a segmentation. We show that these extreme points, or extremal markets, must take a special form: in any extremal market, the monopolist will be indifferent to setting any price in the support of consumers' valuations. Thus, each subset of valuations that includes the given monopoly price generates an extremal market. If the monopolist charges the uniform monopoly price on each extreme segment, we get point A. If he charges the lowest value in the support of each segment (which is also an optimal price, by construction), we get point C; and if he charges the highest value in the support, we get point D. Beyond its welfare implications, this argument also highlights the multiplicity of segmentations that can achieve extreme outcomes, since there are many sets of extremal markets which can be used to decompose a given market.

Thus, we are able to demonstrate that points B, C, and D can be attained. Every point in the shaded triangle in Figure 1 can also be attained, by segmenting a share of the market using extremal markets as in the previous paragraph and segmenting the rest of the market to facilitate perfect price discrimination. Such a segmentation always gives a fixed level of producer surplus between the uniform monopoly profit and perfect price discrimination profit, and the monopolist is indifferent between prices that yield a consumer surplus of zero and prices that maximize social surplus. This gives us a complete characterization of all possible welfare outcomes.

While we focus on welfare implications, we can also completely characterize possible output levels and derive implications for prices. An upper bound on output is the efficient quantity, and this is realized by any segmentation along the efficient frontier. In particular, it is attained in any consumer surplus maximizing segmentation. In such segmentations, prices are always weakly below the uniform monopoly price. We are also able to obtain a tight lower bound on output. Note that in any segmentation, the monopolist must receive at least his uniform monopoly profits, so this profit is a lower bound on social surplus. We say that an allocation is *conditionally efficient* if conditional on the good being sold, it is sold to those with the highest valuations. Such allocations minimize output for a given level of social surplus. In fact, we construct a social surplus minimizing segmentation that results in a conditionally efficient allocation and therefore attains a lower bound on output. In this segmentation, and indeed in any social surplus minimizing segmentation, prices are always weakly higher than the uniform monopoly price.

Though the results described above are for discrete distributions, we are able to prove similar results for continuous demand functions and, more generally, any market that can be described by a Borel measure over valuations. For such distributions, we use a limit argument to establish the existence of segmentations that attain points C and D.

We contribute to a large literature on third degree price discrimination, starting with the classic work of Pigou (1920). This literature examines what happens to prices, quantity, consumer surplus, producer surplus, and social welfare as a market is segmented. Pigou considered the case of two segments with linear demand. In the special case where both segments are served when there is a uniform price, he showed that output does not change under price discrimination. Since different prices are charged in the two segments, this means that some high valuation consumers are replaced by low valuation consumers, and thus social welfare decreases. We can visualize the results of Pigou and other authors in Figure 1. Pigou showed that this particular segmentation resulted in a west-northwest move (i.e., a move from point A to a point below the negative 45° line going through A). A literature since then has focused on identifying sufficient conditions on the shape of demand for social welfare to increase or decrease with price discrimination. A recent paper of Aguirre, Cowan, and Vickers (2010) unifies and extends this literature and, in particular, identifies sufficient conditions for price discrimination to either increase or decrease social welfare (i.e., move above or below the negative 45° line through A). Restricting attention to market segments that have concave profit functions and an additional property ("increasing ratio condition") that they argue is commonly met, they show that welfare decreases if the direct demand in the higher priced market is at least as convex as that in the lower priced market; welfare is higher if prices are not too far apart and the inverse demand function in the lower priced market is locally more convex than that in the higher priced market. They note how their result ties in with an intuition of Robinson (1933): concave demand means that price changes have a small impact on quantity, while convex demand means that prices have a large impact on quantity. If the price rises in a market with concave demand and falls in a market with convex demand, the increase in output in the low-price market will outweigh the decrease in the high price market, and welfare will go up. A recent paper of Cowan (2013) gives sufficient conditions for consumer surplus (and thus total surplus) to increase under third degree price discrimination.

Our paper also gives sufficient conditions for particular welfare impacts of segmentation. However, unlike most of the literature, we allow for segments with non-concave profit functions. Indeed, the segmentations giving rise to extreme points in welfare space (i.e., consumer surplus maximization at point C and social surplus minimization at point D) generally rely on non-concave profit functions within segments. This ensures that the type of local conditions highlighted in the existing literature will not obtain. Our non-local results suggest some very different intuitions. Of course, consumer surplus always increases if prices drop in all markets. We show that for any demand curve, low valuation consumers can be pooled with high valuation consumers in such a way that the producer has an incentive to offer prices below the monopoly price; but if this incentive is made arbitrarily weak, the consumers capture the efficiency gains.

The price discrimination literature also has results on the impact of segmentation on output and prices. With regard to output, the focus of this literature is on identifying when an increase in output is necessary for an increase in welfare, as in Schmalensee (1981) and Varian (1985). Although we do not analyze the question in detail in this paper, a given output level is associated with many different levels of producer, consumer and social surplus. We do provide a sharp characterization of the highest and lowest possible output over all market segmentations. On prices, Nahata, Ostaszewski, and Sahoo (1990) offer examples with non-concave profit functions where third degree price discrimination may lead prices in all market segments to move in the same direction; it may be that all prices increase or all prices decrease. We show that one can create such segmentations for *any* demand curve. In other words, in constructing our critical market segmentations, we show that it is always possible to have all prices rise or all prices fall (although profit functions in the segments cannot all be strictly concave, as shown by Nahata, Ostaszewski, and Sahoo).

If market segmentation is exogenous, one might argue that the segmentations that deliver extremal surpluses are special and might be seen as atypical. But given the amount of information presently being collected on the internet about consumer valuations, it might be argued that there is increasing endogeneity in the market segmentations that arise. To the extent that producers control how information is disseminated, they will have an incentive to gather as much information as possible, ideally engaging in perfect price discrimination. Suppose, however, that an internet intermediary wanted to release its information about consumers to producers for free in order to maximize consumer welfare, say, because of regulatory pressure or as part of a broader business model. Our results describe how such a consumer-minded internet company would choose to structure this information.[1]

Third degree price discrimination is a special case of the classic screening problem, in which a principal is designing a contract for an agent who has private information about the environment. In third degree price discrimination, the gross utility function of the agent is the product of this willingness to pay for the object and the probability of receiving the good. An important observation due to Riley and Zeckhauser (1983) is then that as long as consumers' valuations are linear in the quantity, quality, or probability of getting the object, a posted price is an optimal mechanism. This would no longer be true if valuations are non-linear in quantity as in Maskin and Riley (1984) or in quality as in Mussa and Rosen (1978). In Section 5, we examine the robustness of our result by adding a small amount of concavity to consumers' valuations. In this case, the marginal value of the good to each consumer varies with quantity, as in Maskin and Riley. As such, the monopolist will wish to engage in *second degree price discrimination*, in which consumers are screened through the use of a menu of quantity price bundles. In the non-linear screening environment, welfare bounds can be constructed in a manner analogous to the linear environment of third degree price discrimination. If the principal has no information about the agent's type, then he must offer the same menu to all agents, which yields a uniform menu profit (or producer surplus) for the principal with a corresponding information rent (or consumer surplus) for the agent, leading to a point analogous to A in Figure 1. If the principal was perfectly informed about the agent's type, he could extract all the potential surplus from the relationship, leading to point analogous to B in Figure 1. And, as in the third degree price discrimination problem, there are bounds on surplus pairs for any intermediate segmentation given by a triangle BCD. However, it is not possible in general to find a segmentation that attains

[1]An important subtlety of this story, however, is that this could only be done by randomly allocating consumers with the same valuation to different segments with different prices. Thus, it could be done by a benevolent intermediary who already knew consumers' valuations, but not by one who needed consumers to truthfully report their values.

every point in BCD, and we do not have a characterization of what happens in general screening problems. In an illustrative example with binary types and a continuum of allocations, we do show that as we approach the linear case, the equilibrium surplus pairs converge to the triangle and, in this sense, our main result is robust to small deviations from linearity.

Our work has a methodological connection to two strands of literature. Kamenica and Gentzkow (2011)'s study of "Bayesian persuasion" considers how a sender would choose to transmit information to a receiver, if he could commit to an information revelation strategy before observing his private information. They provide a characterization of such optimal communication strategies as well as applications. If we let the receiver be the producer choosing prices, and let the sender be a planner maximizing some weighted sum of consumer and producer surplus, our problem belongs to the class of problems analyzed by Kamenica and Gentzkow. They show that if one plots the utility of the "sender" as a function of the distribution of the sender's types, his highest attainable utility can be read off from the "concavification" of that function.[2] The concavification arguments are especially powerful in the case of two types. While we do not use concavification arguments in the proof of our main result, we illustrate their use in our two type analysis of second degree price discrimination.

Bergemann and Morris (2013a,b) examine the general question, in strategic many-player settings, of what behavior could arise in an incomplete information game if players observe additional information not known to the analyst. They show that behavior that might arise is equivalent to an incomplete information version of correlated equilibrium termed "Bayes correlated equilibrium", which reduces to the problem of Kamenica and Gentzkow (2011) in the case of one player and no initial information.[3] Using the language of Bergemann and Morris, the present paper considers the game of a producer making take-it-or-leave-it offers to consumers. Here, consumers have a dominant strategy to accept all offers strictly less than their valuation and reject all offers strictly greater than their valuation, and we select for equilibria in which consumers accept offers that make them indifferent. We characterize what could happen for any information structure that players might observe, as long as consumers know their own valuations. Thus, we identify possible payoffs of the producer and consumers in all Bayes correlated equilibria of the price setting game. Thus, our results are a striking

[2]Aumann and Maschler (1995) show that the concavification of the (stage) payoff function represents the limit payoff that an informed player can achieve in a repeated zero sum game with incomplete information. In particular, their Lemma 5.3, the "splitting lemma", derives a partial disclosure strategy on the basis of a concavified payoff function.

[3]In Bergemann and Morris (2013b), these insights were developed in detail in games with a continuum of players, linear-quadratic payoffs and normally distributed uncertainty.

application of the methodologies of Bergemann and Morris (2013a,b) and Kamenica and Gentzkow (2011) to the problem of price discrimination. We also make use of these methodologies as well as results from the present paper in our analysis of what can happen for all information structures in a first-price auction, in Bergemann, Brooks, and Morris (2013a).

We present our model of monopoly price discrimination with discrete valuations in Section 2, with our main results in Section 3. We first give a characterization of the attainable surplus pairs using the extremal segmentations described above. Though the argument for this characterization is non-constructive, we also exhibit some constructive approaches to achieve extreme welfare outcomes. In this Section, we also characterize a tight lower bound on output that can arise under price discrimination. In Section 4, we briefly extend our results to general settings with a continuum of values, so that the demand curve consists of a combination of mass points and densities. The basic economic insights extend to this setting unchanged. In Section 5, we briefly describe how our results change as we move to more general screening environments, where the utility of the buyer and/or the cost of the seller are not linear in quantity, and thus give rise to second degree price discrimination. We conclude in Section 6. Omitted proofs for Sections 3 and 4 are contained in the Appendix, and an Online Appendix contains detailed calculations for the model of Section 5.

2 Model

A monopolist sells a good to a continuum of consumers, each of whom demands one unit. We normalize the total mass of consumers to one and the constant marginal cost of the good to zero.[4] There are K possible values $V \triangleq \{v_1, ..., v_k, ..., v_K\}$, with $v_k \in \mathbb{R}_+$, that the consumers might have, which without loss of generality are increasing in the index k:

$$0 < v_1 < \cdots < v_k < \cdots < v_K.$$

We will extend the analysis to a continuum of valuations in Section 4. A *market* x is a distribution over the K valuations, with the set of all markets being:

$$X \triangleq \Delta(V) = \left\{ x \in \mathbb{R}_+^V \left| \sum_{k=1}^{K} x\left(v_k\right) = 1 \right.\right\}.$$

[4]The assumption of zero marginal cost is without loss of generality: after normalizing $\widehat{v} \triangleq \max\{v - c, 0\}$, all the results apply verbatim to the net social value $\widehat{v}$.

This set can be identified with the $(K-1)$-dimensional simplex, and to simplify notation we will write x_k for $x(v_k)$, which is the proportion of consumers who have valuation v_k. Thus, a market x corresponds to a step demand function, where $\sum_{j=k}^{K} x_j$ is the demand for the good at any price in the interval $(v_{k-1}, v_k]$ (with the convention that $v_0 = 0$).

While we will focus on this interpretation of the model throughout the paper, there is a well-known alternative interpretation that there is a single consumer with unit demand whose valuation is distributed according to the probability distribution x. The analysis is unchanged, and all of our results can be translated into this alternative interpretation.

Throughout the analysis, we hold a given *aggregate market* as fixed and identify it by:

$$x^* \in X.$$

We say that the price v_k is *optimal* for market x if the expected revenue from price v_k satisfies:

$$v_k \sum_{j=k}^{K} x_j \;\geq\; v_i \sum_{j=i}^{K} x_j, \text{ for all } i = 1, ..., K.$$

X_k denotes the set of markets where price v_k is optimal:

$$X_k \triangleq \left\{ x \in X \;\middle|\; v_k \sum_{j=k}^{K} x_j \geq v_i \sum_{j=i}^{K} x_j, \text{ for all } i = 1, ..., K \right\}.$$

Now write $v^* \triangleq v_{i^*}$ for the optimal uniform price for the aggregate market x^*. Thus $x^* \in X^* \triangleq X_{i^*}$. The maximum feasible surplus is:

$$w^* \triangleq \sum_{j=1}^{K} v_j x_j^*,$$

which corresponds to all consumers buying the good. The uniform price producer surplus is then:

$$\pi^* \triangleq v^* \sum_{j=i^*}^{K} x_j^* \;=\; \max_{k \in \{1,...,K\}} v_k \sum_{j=k}^{K} x_j^*;$$

and the uniform price consumer surplus is:

$$u^* \triangleq \sum_{j=i^*}^{K} (v_j - v^*)\, x_j^*.$$

We will use a simple example to illustrate many of the results to follow.

Example 1 (Three Values with Uniform Probability).
There are three valuations, $V = \{1, 2, 3\}$, which arise in equal proportions. Thus, $K = 3$, $v_k = k$, and $x^* = (1/3, 1/3, 1/3)$. The feasible social surplus is $w^* = (1/3)(1 + 2 + 3) = 2$. The uniform monopoly price is $v^* = i^* = 2$. Under the uniform monopoly price, profit is $\pi^* = (2/3) \times 2 = 4/3$ and consumer surplus is $u^* = (1/3)(3 - 2) + (1/3)(2 - 2) = 1/3$.

We can graphically depict markets consisting of three possible valuations using a classic visualization (see Mas-Colell, Whinston, and Green, 1995, page 169). A market is a point in the two-dimensional probability simplex, which forms a triangular plane in three-dimensional Euclidean space as depicted in the left panel of Figure 2. This set can be projected onto two-dimensional Euclidean space as an equilateral triangle, as in the right panel of Figure 2, and we will use this equilateral triangle projection in our subsequent graphical depictions of the example. The extreme points denoted by $x^{\{k\}}$ correspond to markets which put mass 1 on valuation k, and each point in the triangle can be decomposed as a weighted sum of the three vertices with weights corresponding to the respective proportions. Thus, the market $x = (x_1, x_2, x_3)$ is represented as the point $x_1 \cdot x^{\{1\}} + x_2 \cdot x^{\{2\}} + x_3 \cdot x^{\{3\}}$. For example, the uniform market $x^* = (1/3, 1/3, 1/3)$ just discussed is the point equidistant from the three vertices, since it contains equal shares of each valuation.

In the right panel, we have further divided the simplex into the three regions X_1, X_2, and X_3 where prices 1, 2 and 3 are respectively optimal. Note that the restriction that revenue from price v_k is greater than revenue from price v_i is a linear restriction on markets, and thus the region X_1, for example, is the intersection of the region in which price 1 is better than 2 and the region where price 1 is better than 3. Note also that the linear projection from the simplex in $\mathbb{R}^3$ onto the equilateral triangle in $\mathbb{R}^2$ maps linear constraints on markets into half-planes, so that each region X_k is the intersection of the triangle with two half-planes. A price of 2 is strictly optimal for the uniform market x^*, which therefore lies on the interior of X_2.

A *segmentation* is a division of the aggregate market into different markets. Thus, a segmentation σ is a simple probability distribution on X, with the interpretation that $\sigma(x)$ is the proportion of the population in market x. A segmentation can be viewed as a two-stage lottery on outcomes in V whose reduced lottery is x^*. Writing supp for the support of a distribution, the set of possible segmentations is:

$$\Sigma = \left\{ \sigma \in \Delta(X) \,\middle|\, \sum_{x \in \operatorname{supp} \sigma} \sigma(x) \cdot x = x^*, \ \ |\operatorname{supp} \sigma| < \infty \right\}.$$

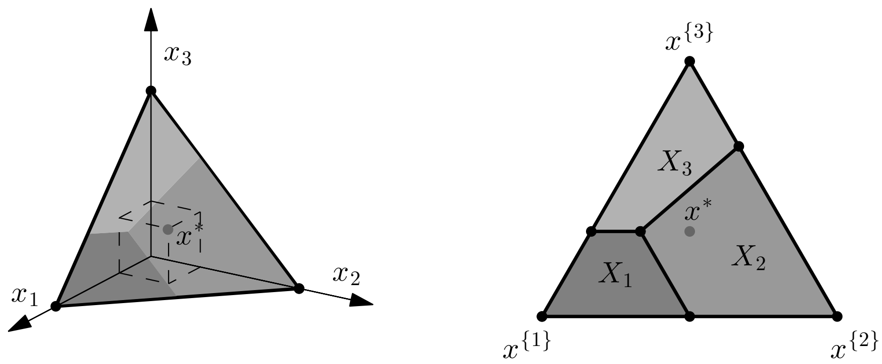

Figure 2: The simplex of markets with $v_k \in \{1, 2, 3\}$.

We restrict attention to finitely many segments so that $|\operatorname{supp} \sigma| < \infty$. This is without loss of generality in the present environment with finitely many valuations, in that finite segmentations will suffice to prove tightness of our bounds on welfare outcomes.

A *pricing rule* for a segmentation σ specifies a distribution over prices for each market in the support of σ:

$$\phi : \operatorname{supp} \sigma \to \Delta(V).$$

We will write $\phi_k(x)$ for the probability of charging price v_k in market x. A pricing rule is *optimal* if, for each x, $v_k \in \operatorname{supp} \phi(x)$ implies $x \in X_k$, i.e., all prices charged with positive probability in market x must be profit maximizing for market x.

An example of a segmentation and an associated optimal pricing rule is given by the case of perfect (or first degree) price discrimination. In this case, there are at least as many segments as possible valuations, and each segment contains consumers of a single valuation. The optimal pricing rule charges the unique valuation that appears in the segment. For Example 1, perfect price discrimination consists of three market segments with three associated prices, as illustrated in the table below:

Segment	x_1	x_2	x_3	$\sigma(x)$	$\operatorname{supp} \phi(x)$
$x^{\{1\}}$	1	0	0	$\frac{1}{3}$	$\{1\}$
$x^{\{2\}}$	0	1	0	$\frac{1}{3}$	$\{2\}$
$x^{\{3\}}$	0	0	1	$\frac{1}{3}$	$\{3\}$
x^*	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1	

(1)

This segmentation can be visualized as simply saying that the point $(1/3, 1/3, 1/3)$ can be decomposed as the convex combination of the vertices $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$.

Given a segmentation σ and pricing rule ϕ, consumer surplus is:

$$\sum_{x \in \operatorname{supp} \sigma} \sigma(x) \sum_{k=1}^{K} \phi_k(x) \sum_{j=k}^{K} (v_j - v_k) x_j;$$

producer surplus is:

$$\sum_{x \in \operatorname{supp} \sigma} \sigma(x) \sum_{k=1}^{K} \phi_k(x) v_k \sum_{j=k}^{K} x_j;$$

and the total surplus is:

$$\sum_{x \in \operatorname{supp} \sigma} \sigma(x) \sum_{k=1}^{K} \phi_k(x) \sum_{j=k}^{K} v_j x_j.$$

3 The Limits of Discrimination

We now turn to the characterization of the set of welfare outcomes which can arise under third degree price discrimination. We will demonstrate that the welfare bounds described in the Introduction are tight, using the special geometry that comes from grouping the markets by corresponding optimal prices. This geometry is the subject of Section 3.1, and we use it to prove our main result in Section 3.2. Though the argument we present is non-constructive, there are in fact many ways of constructing segmentations that achieve the bounds, and we give examples of such constructions in Section 3.3. We will also provide a tight characterization of limits on output in Section 3.4.

3.1 Extremal Markets

Our first result is a linear algebraic characterization of the set X_k of markets where price v_k is optimal. For this purpose, we denote by $\mathcal{V}$ the set of non-empty subsets of $V = \{v_1, ..., v_K\}$. We write $\mathcal{V}_k$ for the set of subsets of V that contain v_k, and $\mathcal{V}^* = \mathcal{V}_{i^*}$ denotes the subsets that contain the aggregate monopoly price v^*. For every subset $S \in \mathcal{V}$, we define a market $x^S \in X$, with the properties that: (i) no consumer has a valuation outside the set S and (ii) the monopolist is indifferent between charging any price inside the set S. Thus, for $S \in \mathcal{V}_k$, the market x^S is then uniquely identified by the *indifference conditions* that if $v_i \in S \setminus \{v_k\}$, then:

$$v_i \sum_{j=i}^{K} x_j^S = v_k \sum_{j=k}^{K} x_j^S; \tag{2}$$

and by the *support condition* that:

$$\sum_{\{j|v_j \in S\}} x_j^S = 1. \tag{3}$$

The above conditions, (2) and (3), represent $|S|$ equations in $|S|$ unknowns, which have a unique solution described below. Writing $\min S$ for the smallest element of S, (2) implies that profits from any price in the support must be $\min S$ and thus we must have:

$$v_i \sum_{j=i}^{K} x_j^S = \min S$$

for all $v_i \in S$. Writing $\mu(v_i, S)$ for the smallest element of S which is strictly greater than v_i, we must have:

$$x_i^S \triangleq \begin{cases} 0, & \text{if} \quad v_i \notin S; \\ \min S \left(\frac{1}{v_i} - \frac{1}{\mu(v_i,S)}\right), & \text{if} \quad v_i \in S \text{ and } v_i \neq \max S; \\ \frac{\min S}{\max S}, & \text{if} \quad v_i = \max S. \end{cases}$$

The indifference conditions (2) can equivalently be expressed in terms of the discrete version of the virtual utility. The fact that the seller is indifferent between selling at any valuation $v_i \in S$ and the highest valuation in S, namely $\max S$, simply means that the virtual utility of any valuation $v_i \in S$ except $\max S$ has to be equal to zero:

$$v_i \in S \backslash \{\max S\} \Leftrightarrow v_i - (\mu(v_i, S) - v_i) \frac{1 - \sum_{k=1}^{i} x_k^S}{x_i^S} = 0.$$

A remarkable and useful property of the set X_k is that every $x \in X_k$ can be represented as a convex combinations of the markets $\{x^S\}_{S \in \mathcal{V}_k}$. For this reason, we will refer to any market of the form x^S for some $S \in \mathcal{V}$ as an *extremal market*. In particular, the extreme points $x^{\{k\}}$ that we introduced in the context of Example 1 are extremal markets, namely those with a support set S that consists of a singleton, $S = \{v_k\} = \{k\}$.

Lemma 1 (Extremal Markets)**.**
X_k is equal to the convex hull of $\{x^S\}_{S \in \mathcal{V}_k}$.

Proof. The inclusion of the convex hull of $\{x^S\}_{S \in \mathcal{V}_k}$ in X_k is immediate, since by definition $x^S \in X_k$ for any $S \in \mathcal{V}_k$, and X_k is convex, being the intersection of the convex simplex and the half spaces in which price v_k is better than price v_i for all $i \neq k$.

Moreover, X_k is finite-dimensional and compact, as it is the intersection of closed sets with the compact simplex. Thus, by the Minkowski-Caratheodory Theorem (Simon, 2011, Theorem 8.11), X_k

is equal to the convex hull of its extreme points. We will show that every extreme point of X_k is equal to x^S for some $S \in \mathcal{V}_k$. First, observe that if v_i is an optimal price for market x, then $x_i > 0$. Otherwise the monopolist would want to deviate to a higher price if $\sum_{j=i+1}^{K} x_j > 0$ or a lower price if this quantity is zero, either of which contradicts the optimality of v_i.

Now, the set X_k is characterized by the linear constraints that for any $x \in X_k$:

$$\sum_{j=1}^{K} x_j = 1;$$

the non-negativity constraints:

$$x_i \geq 0, \text{ for all } i;$$

and the optimality (of price v_k) constraints:

$$v_k \sum_{j=k}^{K} x_j \geq v_i \sum_{j=i}^{K} x_j \text{ for } i \neq k.$$

Any extreme point of X_k must lie at the intersection of at least K of these constraints (see Simon, 2011, Proposition 15.2). One binding constraint is always $\sum_{i=1}^{K} x_i = 1$, and since v_k is an optimal price, the non-negativity constraint $x_k \geq 0$ is always slack. Thus, there must be at least $K - 1$ binding optimality and non-negativity constraints for $i \neq k$.

But as we have argued, we cannot have both the optimality and non-negativity constraints bind for a given i, so for each $i \neq k$ precisely one of these is binding. This profile of constraints defines x^S, where S is the set valuations for which the optimality constraint binds. $\square$

The following is an alternative and intuitive explanation as to why any $\widehat{x} \in X_k \setminus \left\{x^S | \, S \in \mathcal{V}_k\right\}$ cannot be an extreme point of X_k. Let $\widehat{S}$ be the support of $\widehat{x}$ and consider moving from $\widehat{x}$ either towards $x^{\widehat{S}}$ or in the opposite direction. Price v_k will continue to be optimal, because complete indifference at $x^{\widehat{S}}$ to all prices in the support means moving in either direction will not change optimal prices. Also, for small perturbations, we will remain in the simplex, since $\widehat{x}$ and $x^{\widehat{S}}$ have the same support by construction. Since we can move in opposite directions and remain within X_k, it follows that $\widehat{x}$ is not an extreme point of X_k.

We illustrate in Figure 3 the extremal markets for X_2 in the probability simplex of Example 1. Since the uniform monopoly price was $v^* = 2$, the extremal markets x^S corresponding to $S \in \mathcal{V}^*$ are $x^{\{2\}}$, $x^{\{1,2\}}$, $x^{\{2,3\}}$, and $x^{\{1,2,3\}}$. We refer to a segmentation consisting only of extremal markets as an *extremal segmentation* and a segmentation consisting only of extremal markets in X^* as a *uniform*

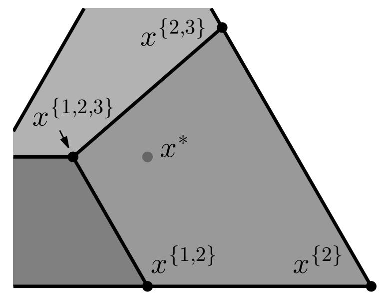

Figure 3: Uniform profit preserving extremal markets for $x^* = (1/3, 1/3, 1/3)$.

profit preserving extremal segmentation. It is a direct consequence of Lemma 1 that a uniform profit preserving extremal segmentation exists.

An example of a uniform profit preserving extremal segmentation for Example 1 is given below.[5]

Segment	x_1	x_2	x_3	$\sigma(x)$
$x^{\{1,2,3\}}$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{2}{3}$
$x^{\{2,3\}}$	0	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{6}$
$x^{\{1\}}$	0	1	0	$\frac{1}{6}$
x^*	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

(4)

An example of an extremal segmentation that is not uniform profit preserving is the perfect price discriminating segmentation described in (1), since this necessarily uses markets $x^{\{v_k\}}$ with $v_k \neq v^*$, which are not in X^*.

3.2 Limits of Welfare

For a given market x, we define the minimum pricing rule $\underline{\phi}(x)$ to deterministically charge $\min \operatorname{supp} x$ and, similarly, we define the maximum pricing rule $\overline{\phi}(x)$ to deterministically charge $\max \operatorname{supp} x$. The minimum pricing rule always implies an efficient allocation in the market x and the maximum pricing rule always implies an allocation in the market x where there is zero consumer surplus. When combined with extremal segmentations, the minimum and maximum pricing rules are especially powerful.

[5]This segmentation is the output of the "greedy" algorithm we will describe in Section 3.3.

Proposition 1 (Extremal Segmentations)**.**
In every extremal segmentation, minimum and maximum pricing rules are optimal. Total surplus is w^* *under the minimum pricing rule, and consumer surplus is zero under the maximum pricing rule. If the extremal segmentation is uniform profit preserving, then producer surplus is* π^* *under every optimal pricing rule, and consumer surplus is* $w^* - \pi^*$ *under the minimum pricing rule.*

Proof. By construction of the extremal markets, any price in S is an optimal price in market x^S. This implies that minimum and maximum pricing rules are both optimal. Under the minimum pricing rule, all consumers purchase the good, so the efficient total surplus is attained. Consumer surplus is always zero under the maximum pricing rule because consumers who purchase pay exactly their value. If the extremal segmentation is uniform profit preserving, setting the price equal to v^* in every segment is optimal, so producer surplus must be exactly π^* under any optimal pricing rule. Combining this with the fact that total surplus is w^* under the minimum pricing rule, we conclude that consumer surplus is $w^* - \pi^*$. □

This Proposition implies that with a uniform profit preserving extremal segmentation, aggregate consumer surplus must be weakly greater under the minimum pricing rule and weakly lower (in particular zero) under the maximum pricing rule. In fact, the same predictions hold conditional on each possible valuation of the consumer. With the minimum pricing rule $\underline{\phi}(x)$, we observe that all efficient trades are realized (as opposed to only those with a value equal to or greater than the uniform monopoly price), and by construction of the minimum pricing rule $\underline{\phi}(x)$, all sales are realized at prices below or equal to v^*. Thus, conditional on each valuation, consumer surplus must increase. As for the maximum pricing rule $\overline{\phi}(x)$, only the buyer with the highest value in the segment x purchases the product but has to pay exactly his valuation. Hence, the expected net utility conditional on a purchase is zero, and so is the expected net utility without a purchase. Thus, all valuation types are weakly worse off relative to the uniform price in the aggregate market.

By construction, every extremal segmentation induces indifference on the part of the monopolist. This is the property that allows a single segmentation to accomplish either the social surplus minimizing or the consumer surplus maximizing allocation. But clearly, the indifference of the monopolist could be turned into a strict preference for a specific allocation by a standard perturbation argument. For example, to realize the consumer surplus maximizing allocation, it would suffice to increase in each segment the proportion of the lowest value relative to the higher values by an arbitrarily small amount, which entails an arbitrarily small transfer of surplus from consumers to the monopolist.

Combining the previous analyses, we have our main result on the welfare limits of price discrimination.

Theorem 1 (Surplus Triangle)**.**
There exists a segmentation and optimal pricing rule with consumer surplus u and producer surplus π if and only if $u \geq 0$, $\pi \geq \pi^*$ and $u + \pi \leq w^*$.

Proof. First we argue necessity. That consumer surplus must be non-negative and that total surplus is bounded above by w^* follows directly from the definitions. For producer surplus, a price offered under an optimal pricing rule must generate weakly greater revenue than would v^*. Summing this inequality over all markets in the segmentation and over all prices induced by the rule yields the desired result:

$$\begin{aligned}
\sum_{x \in \operatorname{supp} \sigma} \sigma(x) \sum_{k=1}^{K} \phi_k(x) v_k \sum_{j=k}^{K} x_j &\geq \sum_{x \in \operatorname{supp} \sigma} \sigma(x) \sum_{k=1}^{K} \phi_k(x) v^* \sum_{j=i^*}^{K} x_j \\
&= v^* \sum_{j=i^*}^{K} x_j^* \\
&= \pi^*.
\end{aligned}$$

For sufficiency, a direct consequence of Lemma 1 is that a uniform profit preserving extremal segmentation σ always exists. By Proposition 1, the minimum and maximum pricing rules under this segmentation achieve the surplus pairs $(w^* - \pi^*, \pi^*)$ and $(0, \pi^*)$ respectively. The segmentation

$$\sigma'(x) = \begin{cases} x_k^*, & \text{if} \quad x = x^{\{v_k\}} \\ 0, & \text{if} \quad \text{otherwise} \end{cases},$$

together with any optimal pricing rule, attains the surplus pair $(0, w^*)$, in which the seller receives the entire surplus. It follows that the three vertices of the surplus triangle can be attained. But now any point in the surplus triangle can be written as a convex combination:

$$\alpha \cdot (0, w^*) + (1 - \alpha) \cdot [\beta \cdot (w^* - \pi^*, \pi^*) + (1 - \beta) \cdot (0, \pi^*)],$$

with $\alpha, \beta \in [0, 1]$. The extremal segmentation:

$$\sigma''(x) = \alpha \sigma'(x) + (1 - \alpha) \sigma(x)$$

together with the optimal pricing rule:

$$\phi_k(x) = \beta \underline{\phi}_k(x) + (1 - \beta) \overline{\phi}_k(x)$$

achieves the desired welfare outcome. □

One implication of the above characterization of the surplus triangle is that the seller cannot be driven below his uniform monopoly profits under any segmentation. We can understand any possible segmentation as additional, possibly noisy, information, that the seller receives beyond the prior distribution. With this interpretation, the fact that the lower bound is given by the uniform monopoly profit is simply a reflection of the general result that information in the sense of Blackwell (1951, 1953) has positive value in single-person decision problems. By contrast, in oligopolistic (and more generally strategic) settings, additional information can lower the profits of the seller below the level obtained with common prior information only (see, e.g., Bergemann and Morris, 2013b).

3.3 Constructive Approaches and Direct Segmentations

Our results thus far establish the existence of segmentations that achieve extreme welfare outcomes, based on the fact that any market can be decomposed as a convex combination of extremal markets in X^*. In general, there will be many such segmentations. One reason is that there may be many subsets of extremal markets in X^* whose convex hulls contain x^*, and therefore many uniform profit preserving extremal segmentations with different supports. A second reason is that extremal segmentations are just one kind of segmentation; welfare bounds can also be attained with segments that are not extremal. We will briefly describe two constructive algorithms in order to give a sense of this multiplicity, to give some intuition for what critical segmentations will end up looking like, and to make some additional observations about the number of segments required. The first algorithm segments the aggregate market by means of a "greedy" procedure and creates segments which are extremal markets as defined earlier in Section 3.1. In fact, it gives an explicit construction of a uniform profit preserving extremal segmentation, whereas our proof of Theorem 1 only relied on the fact that some such segmentation exists. The second algorithm gives a detailed account of the algorithm described informally in the Introduction. It does not use extremal markets but has a narrower objective, namely to construct a segmentation that maximizes the attainable consumer surplus.

We start with a construction of a uniform profit preserving extremal segmentation through the following "greedy" procedure. First, pack as many consumers as possible into the market $x^{\operatorname{supp} x^*}$, i.e., the extremal market in which the monopolist is indifferent between charging all prices in the support of x^*. At some point, we will run out of mass of some valuation in $\operatorname{supp} x^*$, and define the residual market to be the distribution of all remaining consumers. We then proceed inductively with a segment that puts as much mass as possible on the extremal market corresponding to all remaining valuations in the residual market; and so on. At each step, we eliminate at least one valuation from the residual market, so the process will necessarily terminate after at most K rounds.

More formally, let $X^S = \{x \in X \,|\mathrm{supp}\, x \subseteq S\}$, which is the subset of the simplex in which valuations $v_k \notin S$ are assigned zero probability. For example, in the probability simplex of Figure 2, $X^{\{2,3\}}$ is simply the edge between the vertices $x^{\{2\}}$ and $x^{\{3\}}$ where $x_1 = 0$. The extreme points of $X^S \cap X^*$ are simply those $x^{S'}$ such that $i^* \in S'$ and $S' \subseteq S$. The only extreme point of $X^S \cap X^*$ that is in the relative interior of X^S is x^S, since any market with $x_k = 0$ for some $k \in S$ is on the relative boundary of X^S.[6] Continuing the example, $x^{\{2\}}$ and $x^{\{2,3\}}$ are the extreme points of $X^{\{2,3\}} \cap X_2$, but $x^{\{2,3\}}$ is in the relative interior of $X^{\{2,3\}}$ and $x^{\{2\}}$ is on the relative boundary.

With these observations in hand, we can define the greedy algorithm as the following iterative procedure. At the end of iteration $l \geq 0$, the "residual" market of valuations not yet assigned to a segment is x^l, with $x^0 = x^*$, and the support of this residual is defined to be $S_l = \mathrm{supp}\, x^l$. We now describe what happens at iteration l, taking as inputs $\left(x^{l-1}, S_{l-1}\right)$, the residual and the support from the previous iteration. If $x^{l-1} = x^{S_{l-1}}$, then we define $\alpha^l = 1$ and $x^l = 0$. Otherwise, we find the unique $\widehat{t}$ for which the market:

$$z\left(t\right) \triangleq x^{S_{l-1}} + t \cdot \left(x^{l-1} - x^{S_{l-1}}\right)$$

is on the relative boundary of $X^{S_{l-1}}$, and thus $z_k\left(\widehat{t}\right) = 0$ for some $k \in S_{l-1}$. In effect, we are projecting x^{l-1} onto the relative boundary of $X^{S_{l-1}}$ by moving away from $x^{S_{l-1}}$, and $z\left(\widehat{t}\right)$ is this projection. Note that moving away from $x^{S_{l-1}}$ will never take us out of X^*, since this transformation preserves the set of optimal prices. In particular, for any $v_i \in S_{l-1}$, the loss in revenue from pricing at v_i instead of v^* is:

$$v^* \sum_{j=i^*}^{K} z_j\left(t\right) - v_i \sum_{j=i}^{K} z_j\left(t\right) = t\left(v^* \sum_{j=i^*}^{K} x_j^{l-1} - v_i \sum_{j=i}^{K} x_j^{l-1}\right),$$

which is non-negative as long as $t \geq 0$. Also observe that this transformation preserves the fact that $\sum_{j=1}^{K} z_j\left(t\right) = 1$. Finally, since $x^{l-1} \neq x^{S_{l-1}}$, having $x_i^{l-1} \geq x_i^{S_{l-1}}$ for all i would violate probabilities in x_i^{l-1} summing to one. Therefore, there is at least one i for which $x_i^{l-1} < x_i^{S_{l-1}}$, so that $z_i\left(t\right)$ eventually hits zero. The desired $\widehat{t}$ is $\inf\left\{t \geq 0 | z_i\left(t\right) < 0 \text{ for some } i\right\}$. The next residual x^l is defined to be the

[6]The relative interior of X^S is the set of $x \in X^S$ such that for all $x_1, x_2 \in X^S$, $\{x + \epsilon \cdot z, x - \epsilon \cdot z\} \subseteq X^S$ for ϵ sufficiently small, where $z = x_1 - x_2$. The relative boundary is just the set of points in X^S that are not on the relative interior.

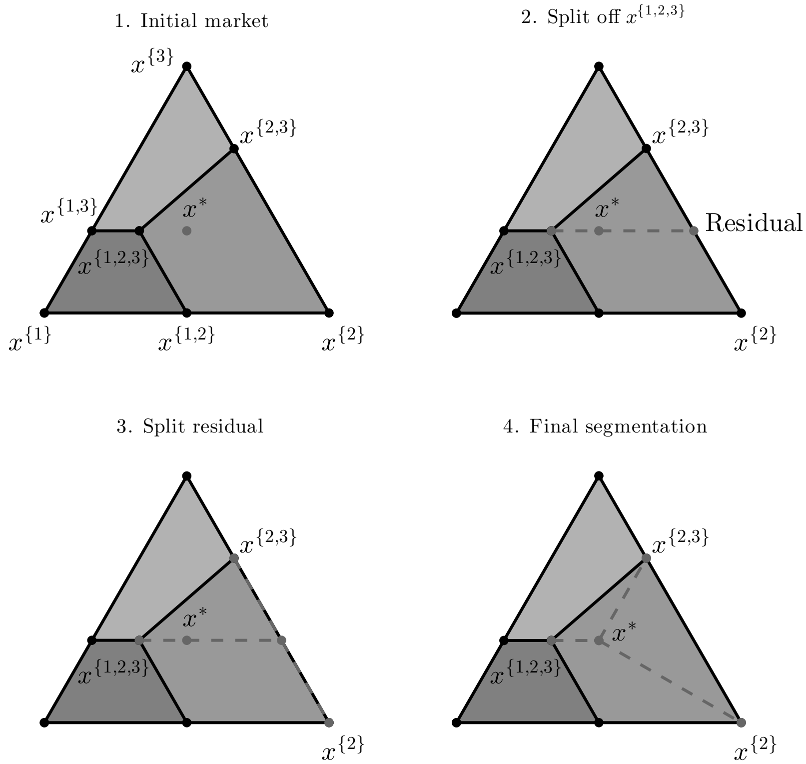

Figure 4: The greedy segmentation for $x^* = (1/3, 1/3, 1/3)$.

projection $z\left(\widehat{t}\right)$, and α^l is defined by:

$$x^{l-1} = \alpha^l \cdot x^{S_{l-1}} + \left(1 - \alpha^l\right) \cdot x^l. \tag{5}$$

Now inductively define $S^l = \operatorname{supp} x^l$, which is a strict subset of S^{l-1} since x^l is on the relative boundary of $X^{S_{l-1}}$.

The inductive hypothesis is that:

$$x^* = \sum_{j=1}^{l} \alpha^j \prod_{i=1}^{j-1} \left(1 - \alpha^i\right) \cdot x^{S_{j-1}} + \prod_{i=1}^{l} \left(1 - \alpha^i\right) \cdot x^l,$$

meaning that the aggregate market is segmented by the x^{S_j} for $j < l$ and the residual market. This property is trivially satisfied for the base case $l = 1$ (with the convention that the empty product is equal to 1). Our choice of α^l in (5) guarantees that if the inductive hypothesis holds at $l - 1$, it will continue to hold at l as well. The algorithm terminates at iteration $L + 1$ when $x^L = x^{S_L}$, which certainly has to be the case when $|S_L| = 1$, and we define the segmentation to have support equal to $\left\{x^{S_l}\right\}_{l=0}^{L}$ with $\sigma\left(x^{S_l}\right) = \alpha^{l+1} \prod_{j=1}^{l} \left(1 - \alpha^j\right)$.

For Example 1, this decomposition is visually depicted in Figure 4. In the first iteration, the aggregate market x^* is projected away from $x^{\{1,2,3\}}$ onto the edge $X^{\{2,3\}}$ at a point in X_2, and therefore between $x^{\{2\}}$ and $x^{\{2,3\}}$. At the second iteration, this residual is further projected away from $x^{\{2,3\}}$ onto $x^{\{2\}}$, at which point the algorithm terminates. This decomposition results in the following segmentation which we previously used to illustrate uniform profit preserving extremal segmentations in (4):

Segment	x_1	x_2	x_3	$\sigma(x)$	$\operatorname{supp} \underline{\phi}(x)$	$\operatorname{supp} \overline{\phi}(x)$
$x^{\{1,2,3\}}$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{2}{3}$	$\{1\}$	$\{3\}$
$x^{\{2,3\}}$	0	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{6}$	$\{2\}$	$\{3\}$
$x^{\{2\}}$	0	1	0	$\frac{1}{6}$	$\{2\}$	$\{2\}$
x^*	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1		

This algorithm incidentally establishes constructively that at most K segments are required to attain points on the bottom of the welfare triangle, and thus at most $2K$ segments to attain all points in the welfare triangle.

Notice that in the greedy segmentation of Example 1, under either the minimum or the maximum pricing rule, there are multiple segments in which the same price is charged. For example, if we focus on maximizing consumer surplus, then the monopolist is to charge price 2 when the segment is $x^{\{2,3\}}$ or when the segment is $x^{\{2\}}$. From the monopolist's point of view, he would also be happy to charge price 2 if we just told him that the market was one of $x^{\{2,3\}}$ and $x^{\{2\}}$, but we did not specify which one. The reason is that price 2 is also optimal in the "merged" market:

$$\frac{\sigma(x^{\{2,3\}})}{\sigma(x^{\{2,3\}})+\sigma(x^{\{2\}})}\cdot x^{\{2,3\}}+\frac{\sigma\left(x^{\{2\}}\right)}{\sigma(x^{\{2,3\}})+\sigma(x^{\{2\}})}\cdot x^{\{2\}}.$$

Given this observation, a natural class of segmentations (and associated pricing rules) are those in which any given price is charged in at most one segment. Formally, we define a *direct segmentation* σ to be one that has support on at most K markets, indexed by $k \in \{1,..,K\}$ such that $x^k \in X_k$. In other words, price v_k is optimal on its corresponding segment x^k. The *direct pricing rule* is the rule that puts probability one on price v_k being charged in market x^k, i.e., $\phi_k(x^k) = 1$. This notation is in contrast to the extremal markets where the upper case superscript S in x^S referred to the support of the distribution. Here, the lower case superscript k in x^k refers to the price v_k charged in the direct segment x^k. By definition, the direct pricing rule is optimal for direct segmentations, and whenever we refer to a direct segmentation in the subsequent discussion, it is assumed that the monopolist will use direct pricing.

Extremal segmentations and direct segmentations are both rich enough classes to achieve any welfare outcome. The reason is that welfare is completely determined by the joint distribution over prices and valuations that is induced by the segmentation and the pricing rule, and both classes of segmentations can achieve any such joint distribution.

Proposition 2 (Extremal and Direct Segmentations)**.**
For any segmentation and optimal pricing rule (σ, ϕ), there exist: (i) an extremal segmentation and an optimal pricing rule (σ', ϕ') and (ii) a direct segmentation σ'' (and associated direct pricing rule ϕ'') that achieve the same joint distribution over valuations and prices. As such, they achieve the same producer surplus, consumer surplus, total surplus, and output.

Proof. To find an extremal segmentation, each market $x \in \operatorname{supp} \sigma$ can itself be decomposed using extremal markets with a segmentation σ_x, using only those indifference sets S which contain $\operatorname{supp} \phi(x)$. The extremal segmentation of (σ, ϕ) is then defined by:

$$\sigma'(x^S) \triangleq \sum_{x \in \operatorname{supp} \sigma} \sigma(x) \sigma_x(x^S),$$

and the corresponding pricing rule is:

$$\phi'_k(x^S) \triangleq \frac{1}{\sigma'(x^S)} \sum_{x \in \operatorname{supp} \sigma} \sigma(x) \sigma_x(x^S) \phi_k(x).$$

Similarly, the direct segmentation σ'' can be defined by:

$$\sigma''(x^k) \triangleq \sum_{x \in \operatorname{supp} \sigma} \sigma(x) \phi_k(x),$$

and therefore:

$$x^k \triangleq \frac{1}{\sigma''(x^k)} \sum_{x \in \operatorname{supp} \sigma} \sigma(x) \phi_k(x) \cdot x$$

yields the corresponding composition of each direct segment x^k. □

As an example, the direct segmentation corresponding to the consumer surplus maximizing greedy extremal segmentation of Example 1 is:

Segment	x_1	x_2	x_3	$\sigma(x)$	$\operatorname{supp}\phi(x)$
$x^1 = x^{\{1,2,3\}}$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{2}{3}$	$\{1\}$
$x^2 = \frac{1}{2}\left(x^{\{2\}} + x^{\{2,3\}}\right)$	0	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\{2\}$
x^*	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1	

where the market for price 3 is degenerate. Note that x^1 is extremal but x^2 is not, being the convex combination of the extremal markets $x^{\{2\}}$ and $x^{\{2,3\}}$. This direct segmentation is visually represented in Figure 4.2 and is realized after the first step of the greedy algorithm, namely after splitting $x^{\{1,2,3\}}$ from the aggregate market. This example illustrates the observation that, while any surplus pair can be achieved by either a direct segmentation or an extremal segmentation, it is generally not possible to attain any surplus pair with a segmentation that is both extremal and direct. To see why, observe that in Example 1, any extremal segmentation must use at least three segments, while any direct segmentation that attains maximum consumer surplus must use at most two segments.

Direct segmentations are a convenient tool for constructing some alternative and intuitive segmentations that attain the welfare bounds. Let us give a formal description of the first segmentation described in the Introduction that attains maximum consumer surplus. For each $k \leq i^*$, let market x^k have the features that: (i) the lowest valuation in the support is v_k and (ii) all values of v_{k+1} and above appear in the same relative proportion as in the aggregate population:

$$x_i^k \triangleq \begin{cases} 0, & \text{if} \quad i < k, \\ 1 - \gamma_k \sum_{j=k+1}^{K} x_j^*, & \text{if} \quad i = k, \\ \gamma_k x_i^*, & \text{if} \quad i > k, \end{cases}$$

where $\gamma_k \in [0,1]$ uniquely solves:

$$v_k \left(x_k^* + \gamma_k \sum_{j=k+1}^{K} x_j^* \right) = \gamma_k v^* \sum_{j=i^*}^{K} x_j^*.$$

By the above equality, both v_k and v^* are optimal prices for segment x^k. We can always construct a segmentation of the aggregate market x^* that uses only $\left\{x^k\right\}_{k=1}^{i^*}$. We establish the construction inductively, letting:

$$\sigma\left(x^1\right) \triangleq \frac{x_1^*}{x_1^1},$$

and:

$$\sigma\left(x^{k}\right) \triangleq \frac{x_{k}^{*}-\sum_{j=1}^{k-1} \sigma\left(x^{j}\right) x_{k}^{j}}{x_{k}^{k}}.$$

We can verify that this segmentation generates maximum consumer surplus by charging in segment x^k the price v_k. The direct pricing rule is optimal and gives rise to an efficient allocation, and because the monopolist is always indifferent to charging v^*, producer surplus is π^*.

Direct segmentations correspond to direct mechanisms in mechanism design. They are minimally informative in the sense that among all information structures which induce a given joint distribution over prices and values, the information structure in the direct segmentation is the least informative according to the ranking of Blackwell (1951). While extremal segmentations have been a key tool in this setting, in many related applications (such as those in Kamenica and Gentzkow (2011) for the single player case and in Bergemann and Morris (2013a,b) for the many player case), it is more convenient to work with direct segmentations.

3.4 Limits of Output

We conclude this Section by applying our methods to the question of how the quantity produced varies across segmentations. The literature has long been interested in the relationship between output, welfare, and discriminatory pricing, as in the classic works of Schmalensee (1981) and Varian (1985). These papers showed that for price discrimination to lead to an increase welfare, output must increase as well. The reason is that with a uniform price, the allocation of the good is *conditionally efficient*, meaning that conditional on being sold, the good is sold to those with the highest valuations. Thus, any allocation that generates greater surplus must involve greater output. We will show that there are simple and tight bounds on the range of output levels that can be attained through market segmentation, with the logic being closely related to that of Schmalensee and Varian: the critical segmentations that generate extreme output levels are those that induce conditionally efficient allocations.

For a segmentation σ and pricing rule ϕ, output is given by:

$$\sum_{x \in \operatorname{supp} \sigma} \sigma(x) \sum_{k=1}^{K} \phi_{k}(x) \sum_{j=k}^{K} x_{j}.$$

An upper bound on output among all segmentations and optimal pricing rules is selling to all consumers, and this bound is achieved by any efficient segmentation.[7] Characterizing the lowest possible output is more subtle. We will first establish a lower bound and then show that it can be attained.

To establish a lower bound on output, recall that the producer must get at least the uniform monopoly profits π^*, which is in turn a lower bound on total surplus that requires some positive output. The smallest output delivering a total surplus of π^* will arise in a conditionally efficient allocation. In our discrete model, this means that there must be a critical valuation $v_{\underline{i}}$ such that the good is always sold to all consumers with valuations above $v_{\underline{i}}$ and never sold to consumers with valuations below $v_{\underline{i}}$. Thus, letting $\underline{i}$ and $\underline{\beta} \in (0,1]$ uniquely solve:

$$v_{\underline{i}}\underline{\beta}x^*_{\underline{i}} + \sum_{j=\underline{i}+1}^{K} v_j x^*_j = \pi^*, \tag{6}$$

we obtain a lower bound $\underline{q}$ given by:

$$\underline{q} \triangleq \underline{\beta}x^*_{\underline{i}} + \sum_{j=\underline{i}+1}^{K} x^*_j.$$

The additional variable $\underline{\beta} \in (0,1]$ describes the proportion of buyers at the threshold value $v_{\underline{i}}$ who must purchase the good to achieve equality in (6) in this discrete setting.

With respect to the earlier Example 1, we have $\underline{q} = \underline{\beta} = 1/2$ and $\underline{i} = i^* = 2$. In fact, the greedy segmentation for Example 1 displayed in (4) in combination with the maximum pricing rule induces a conditionally efficient: all of the consumers with valuation 3 and exactly half of the consumers with valuation 2 purchase the good. However, it is not the case more generally that every uniform profit preserving extremal segmentation delivers a conditionally efficient outcome under the maximum pricing rule, as illustrated in the following example.

Example 2 (Three Values without Uniform Probability)**.**
The setup is the same as in Example 1, except that now the proportion of valuation 1 consumers is $x^*_1 = 3/5$*, and the proportions of valuations 2 and 3 are* $x^*_2 = x^*_3 = 1/5$*. The monopoly price is* 1*,* $\pi^* = 1$*,* $u^* = 3/5$*, and* $w^* = 8/5$*. The minimum output is* $\underline{q} = 2/5$*,* $\underline{i} = 2$*, and* $\underline{\beta} = 1$*.*

[7] If we had a positive constant marginal cost and some consumers had valuations below the marginal cost, then the producer could never be induced to sell to those consumers, so it would still be the case that the efficient output would be an upper bound that was attained.

Two alternative uniform profit preserving extremal segmentations of $x^* = (3/5, 1/5, 1/5)$ are:

Segment	x_1	x_2	x_3	$\sigma(x)$	$\text{supp}\,\overline{\phi}(x)$
Segmentation 1					
$x^{\{1,2,3\}}$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{3}{5}$	$\{3\}$
$x^{\{1,2\}}$	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{5}$	$\{2\}$
$x^{\{1\}}$	1	0	0	$\frac{1}{5}$	$\{1\}$
Segmentation 2					
$x^{\{1,3\}}$	$\frac{2}{3}$	0	$\frac{1}{3}$	$\frac{3}{5}$	$\{3\}$
$x^{\{1,2\}}$	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{2}{5}$	$\{2\}$
x^*	$\frac{3}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	1	

Segmentation 1 is constructed by the greedy algorithm described in the previous section, and is illustrated with dashed lines in the right panel of Figure 5. However, only Segmentation 2, illustrated with light lines, leads to a conditionally efficient allocation with the maximum pricing rule, in which case only prices 2 and 3 are used. In Segmentation 1, a price of 1 will be charged in segment $x^{\{1\}}$ under maximum pricing, thus inducing a conditional inefficiency.

In spite of this apparent complication, it turns out that for any aggregate market, it is always possible to find a uniform profit preserving extremal segmentation that, together with the maximum pricing rule, results in a conditionally efficient outcome. Our approach is analogous to that employed in Section 3.2. We will show that when looking for a uniform profit preserving extremal segmentation, it is without loss of generality to look at a particular subset of extremal markets in X^*. The use of these markets will then imply that the allocation under the maximum pricing rule is conditionally efficient. In particular, we divide the simplex into regions $X_{k,l}$ for $l = k, \ldots, K$, which is the set of markets in X_k for which v_l is the lowest valuation receiving the good in the minimum quantity conditionally efficient outcome. In other words:

$$X_{k,l} \triangleq \left\{ x \in X_k \;\middle|\; \sum_{j=l+1}^{K} v_j x_j \leq v_k \sum_{j=k}^{K} x_j \leq \sum_{j=l}^{K} v_j x_j \right\}.$$

We will refer to these two additional constraints in the definition of $X_{k,l}$ as the lower and upper output constraints (LC and UC, respectively). Note that with $l = K$, the left-hand side of the lower constraint is zero, and with $l = k$, the right-hand side of the upper constraint is necessarily at least π^*, so

$$\bigcup_{l=1}^{K} X_{k,l} = \left\{ x \in X_k \,\middle|\, 0 \leq v_k \sum_{j=k}^{K} x_j \leq \sum_{j=k}^{K} v_j x_j \right\} = X_k.$$

Let $\mathcal{V}_{k,l}$ be the subsets of values:

$$\mathcal{V}_{k,l} \triangleq \left\{ S \in \mathcal{V}_k \,\middle|\, \begin{array}{c} S \cap \{v_l, \ldots, v_K\} \neq \emptyset, \\ |S \cap \{v_{l+1}, \ldots, v_K\}| \leq 1 \end{array} \right\}.$$

We have the following linear algebraic characterization of the sets $X_{k,l}$ in terms of extremal markets with supports in $\mathcal{V}_{k,l}$, which mirrors Lemma 1.

Lemma 2 (Extremal Markets with Output Constraints)**.**
$X_{k,l}$ is the convex hull of $\left\{x^S\right\}_{S \in \mathcal{V}_{k,l}}$.

We can now verify that the output lower bound is attained.

Proposition 3 (Quantity Minimizing Segmentation)**.**
For every market, there exists a uniform profit preserving extremal segmentation such that the allocation under the maximum pricing rule is conditionally efficient. As a result, producer surplus is π^*, consumer surplus is 0, and output is $\underline{q}$.

Proof. By definition, $x^* \in X_{i^*,l}$ for $l = \underline{i}$. As such, x^* can be written as a convex combination of extreme points of $X_{i^*,l}$. For every such market, $S \cap \{v_l, \ldots, v_K\} \neq \emptyset$ implies only consumers with valuations weakly greater than v_l receive the good, and $|S \cap \{v_{l+1}, \ldots, v_K\}| \leq 1$ implies that all consumers with valuations strictly greater than v_l purchase the good. For if not, there must be some market in which consumers with valuation $v_i > v_l$ have strictly positive mass but the price charged is $v_{i'} > v_i$. But this can only happen if $\{v_i, v_{i'}\} \subseteq S$, a contradiction. As a result, the allocation is conditionally efficient, but because the segmentation is uniform profit preserving extremal, producer surplus is π^*, and under the maximum pricing rule consumer surplus is 0. Hence, output is $\underline{q}$. $\square$

For aggregate markets with three possible valuations, the geometry is quite simple when the optimal uniform price is $v^* = 2$ or $v^* = 3$. In those cases, every extremal segmentation results in a conditionally efficient outcome whenever the maximum pricing rule is used: for $v^* = 3$, the maximum pricing rule always induces a price of 3, and when $v^* = 2$, one can verify from the definitions that $X_{2,2} = X_2$. However, when $v^* = 1$, as in Example 2, then the sets $X_{1,1}$ and $X_{1,2}$ have disjoint and

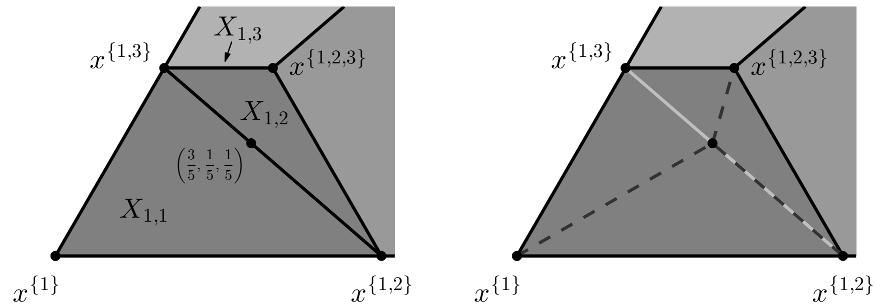

Figure 5: Multiple extremal segmentations.

non-empty interiors. $X_{1,3}$ is just the line segment connecting $x^{\{1,3\}}$ and $x^{\{1,2,3\}}$. This is illustrated in the left panel of Figure 5. Here, the market $x^* = (3/5, 1/5, 1/5)$ lies on the boundary of $X_{1,1}$ and $X_{1,2}$, and it also lies in the convex hull of $\left\{x^{\{1\}}, x^{\{1,2\}}, x^{\{1,2,3\}}\right\}$. The latter sets of extremal markets appear in the segmentation generated by the greedy algorithm.

4 A Continuum of Valuations

Until now, we have considered markets that have a finite support of valuations. The finite structure has allowed us to make simple geometric arguments to characterize the limits of price discrimination. Nonetheless, our results generalize in a straightforward manner to environments with infinitely many valuations. In this Section, we give a simple convergence argument showing why this is the case, and we report some examples of critical segmentations for continuous demand curves.

For the present analysis, we redefine a market to be an element $x \in X = \Delta\left([0, \overline{v}]\right)$, which is the set of Borel probability measures on the interval $[0, \overline{v}]$, and we endow the set X with the weak-* topology. We will write $x\left(Y\right)$ for the measure of a set $Y \in \mathcal{B}\left([0, \overline{v}]\right)$, which is the collection of Borel subsets of $[0, \overline{v}]$. As before, we fix an aggregate market $x^* \in X$, and let v^* denote a uniform monopoly price that solves:

$$v^* \in \underset{v \in [0,\overline{v}]}{\arg\max}\, v\, x^*\left([v, \overline{v}]\right).$$

Note that for any Borel measure, such a maximizer exists due to the fact that $x^*\left([v, \overline{v}]\right)$ is monotonically decreasing and continuous from the left.

The set X_v is defined to be the set of markets in which v is a maximizer of $v'\, x\left([v', \overline{v}]\right)$. We let $\overline{X}_v$ denote the set of markets in X_v such that the seller is indifferent between setting any price in the support, i.e.:

$$\overline{X}_v = \{x \in X_v \,|v' \, x\left([v', \overline{v}]\right) = \min \operatorname{supp} x, \text{ for all } v' \in \operatorname{supp} x\}.$$

We write $\widehat{X}_v$ for the subset of $\overline{X}_v$ with finite support. A preliminary result asserts that a convergent sequence of measures in $\widehat{X}_v$ must converge to an element of $\overline{X}_v$.

Lemma 3 (Closure)**.**
cl $\widehat{X}_v \subseteq \overline{X}_v$.

In fact, the closure of $\widehat{X}_v$ is equal to $\overline{X}_v$, but the weaker property is sufficient for our goals. Lemma 3 should not be viewed as a continuous analogue of Lemma 1, but rather as the "glue" that binds the discrete characterization of extremal markets to the continuous characterization. Extremal markets in the continuous case are precisely the elements of $\overline{X}_v$ for some v, but in order to find an extremal segmentation in the proof of Proposition 1B, which appears in the Appendix, we will take limits of extremal segmentations of finite approximations to x^*. Convergence of this sequence of segmentations is implied by Lemma 3.

To that end, we redefine a *segmentation* of the market x^* to be an element $\sigma \in \Sigma$, where:

$$\Sigma = \left\{\sigma \in \Delta\left(X\right) \left| \int_{x \in X} x\left(Y\right) \sigma\left(dx\right) = x^*\left(Y\right), \text{ for all } Y \in \mathcal{B}\left([0, \overline{v}]\right)\right.\right\}.$$

A segmentation σ is *uniform profit preserving extremal* if its support is contained in $\overline{X}_{v^*}$. A *pricing rule* is a mapping $\phi : \operatorname{supp} \sigma \to X$, and the pricing rule is *optimal* if for all $x \in \operatorname{supp} \sigma$, $\operatorname{supp} \phi\left(x\right) \subseteq \arg\max_{v \in [0,\overline{v}]} v\, x\left([v, \overline{v}]\right)$. The *minimum pricing rule* and *maximum pricing rule* put probability one on the minimum and maximum of $\operatorname{supp} x$ for all $x \in \operatorname{supp} \sigma$, respectively. Consumer surplus is:

$$\int_{x \in X} \int_{v \in [0,\overline{v}]} \int_{v' \in [v,\overline{v}]} \left(v' - v\right) \, x\left(dv'\right) \phi\left(dv\right) \sigma\left(dx\right);$$

producer surplus is:

$$\int_{x \in X} \int_{v \in [0,\overline{v}]} v\, x\left([v, \overline{v}]\right) \phi\left(dv\right) \sigma\left(dx\right);$$

and total surplus is:

$$\int_{x \in X} \int_{v \in [0,\overline{v}]} \int_{v' \in [v,\overline{v}]} v'\, x\left(dv'\right) \phi\left(dv\right) \sigma\left(dx\right).$$

We can then re-establish the earlier Proposition 1 and Theorem 1 for the environment with a Borel measurable aggregate market.

Proposition 1B (Extremal Segmentations with a Continuum of Values)**.**
In every extremal segmentation, minimum and maximum pricing rules are optimal. Total surplus is

w^* under the minimum pricing rule, and consumer surplus is zero under the maximum pricing rule. If the extremal segmentation is uniform profit preserving, then producer surplus is π^* under every optimal pricing rule, and consumer surplus is $w^* - \pi^*$ under the minimum pricing rule.

Combining results, we have the following theorem.

Theorem 1B (Surplus Triangle with a Continuum of Values)**.**
There exists a segmentation and optimal pricing rule with consumer surplus u and producer surplus π if and only if $u \geq 0$, $\pi \geq \pi^*$ and $u + \pi \leq w^*$.

We illustrate the continuum characterization with the following example, for which we can derive explicit segmentations with a convex support for all segments.

Example 3 (Unit Interval with Uniform Density)**.**
The valuations of the consumers are uniformly distributed between 0 and 1, so that $x^*([v, 1]) = 1 - v$ for all $v \in [0, 1]$. The uniform monopoly price is $v^* = 1/2$, uniform monopoly profits are $\pi^* = 1/4$, and the efficient surplus is $w^* = 1/2$.

We will construct a uniform profit preserving extremal segmentation σ of this x^* in which there is a uniform distribution of market segments x_p for $p \in [0, 1/2]$. Each segment x_p has support of the form $[p, z(p)]$, and is defined by:

$$x_p([v, 1]) = \begin{cases} 1, & \text{if} \quad v \leq p, \\ \frac{p}{v}, & \text{if} \quad p < v \leq z(p), \\ 0, & \text{if} \quad v > z(p), \end{cases}$$

and the upper boundary point is given by:

$$z(p) = \frac{1 + \sqrt{1 - 4p^2}}{2},$$

which is monotonically decreasing and has range $[1/2, 1]$. Thus, the support sets of the segments can be ordered by the strong set order. By construction, all of the segments are in $X_{1/2}$, and in fact the segmentation arises as the solution to the continuous version of the greedy algorithm constructed in Section 3.3.

Let us briefly verify that σ is in fact a segmentation of the aggregate market. It is sufficient to check that the density of a valuation v integrates to one. For $v \in [0, 1/2]$, the density in market x_p

is p/v^2 when $p \leq v$ and zero otherwise, so the aggregate density is:

$$\int_{p=0}^{v} 2\frac{p}{v^2}dp = \left.\frac{p^2}{v^2}\right|_{p=0}^{v} = 1.$$

If $v \in [1/2, 1]$, the density in market x_p is p/v^2 when $z(p) > v$, and there is a conditional mass point of size p/v in the market x_p such that $z(p) = v$, which is when $p = \sqrt{v(1-v)}$. Note that the probability that the maximum of the support of x_p is less than w is $1 - 2\sqrt{w(1-w)}$, so the density at v is $(2v-1)/\sqrt{v(1-v)}$. Therefore the aggregate density for every valuation v is equal to:

$$\frac{2v-1}{\sqrt{v(1-v)}}\frac{\sqrt{v(1-v)}}{v} + \int_{p=0}^{\sqrt{v(1-v)}} 2\frac{p}{v^2}dp = \frac{2v-1}{v} + \frac{1-v}{v} = 1.$$

We conclude that σ does in fact segment the aggregate market, as it preserves the aggregate density.

The proof of Theorem 1B establishes the existence of uniform profit preserving extremal segmentations for general Borel measurable distributions. In Bergemann, Brooks, and Morris (2013b), we establish a related existence result for direct segmentations, stated as Theorem 2 in that paper. In addition, there we show that when we narrow the analysis to aggregate markets with differentiable distribution functions, we can explicitly construct direct segmentations that achieve the extreme welfare outcomes as solutions of differential equations. The resulting segmentations, given in Bergemann, Brooks, and Morris (2013b) as Propositions 3 and 4, mirror those in the finite environment (see our current Proposition 2).

We will illustrate these results with examples of direct segmentations for the uniform environment of Example 3. The consumer surplus maximizing segmentation, as derived there in Proposition 3, leads to an associated distribution function of prices $\overline{H}(p)$ given by:

$$\overline{H}(p) = 1 - \frac{1-p}{1-2p}e^{-\frac{2p}{1-2p}}, \text{ for } p \in \left[0, \frac{1}{2}\right].$$

By contrast, the segmentation of the consumers in the total surplus and output minimizing allocation as described there by Proposition 4 leads to a distribution function of prices given by:

$$\underline{H}(p) = 2p^2 - 1, \text{ for } p \in \left[\frac{1}{\sqrt{2}}, 1\right].$$

The distributions of prices induced by these distinct direct segmentations are displayed in the left panel of Figure 6, where the upper curve represents the consumer surplus maximizing price distribution, and the lower curve represents the total surplus minimizing price distribution.

The consumer surplus maximizing and total surplus minimizing distributions represent optimal pricing policies for distinct segmentations of the same aggregate market. And even though they

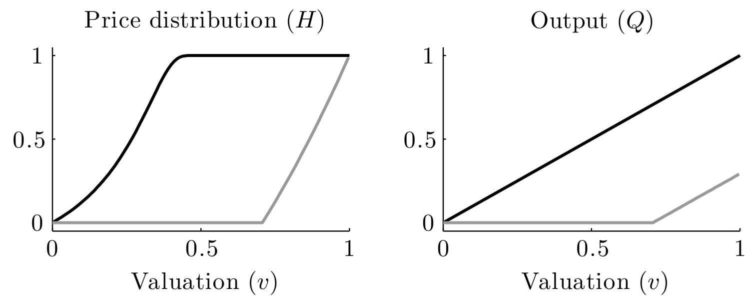

Figure 6: Price and output distribution in direct total surplus minimizing and consumer surplus maximizing segmentations.

share the same aggregate market, the support sets of prices do not overlap.[8] These distinct price distributions also lead to very different allocations. The surplus maximizing pricing policy generates all efficient sales, and hence the cumulative distribution of sales, $\overline{Q}(v) = v$, exactly replicates the aggregate distribution of consumers' valuations. By contrast, the surplus minimizing distribution truncates sales for values v below $1/\sqrt{2}$. As we described in Proposition 2, the allocation is conditionally efficient, and hence $\underline{Q}(v) = v - 1/\sqrt{2}$ for $v \in [1/\sqrt{2}, 1]$, and zero for $v < 1/\sqrt{2}$. These different patterns of sales are displayed in the right panel of Figure 6, where the upper curve represents the consumer surplus maximizing output distribution, and the lower curve represents the total surplus minimizing output distribution.

5 Beyond the Linear Case

We have thus far established the limits of price discrimination in the canonical model of monopoly pricing. The monopoly problem with unit demand may be viewed as special case of a more general class of screening problems, as considered in the seminal papers by Mussa and Rosen (1978) and Maskin and Riley (1984) and often referred to as second degree price discrimination. In these models, the utility of the consumer, or the cost of the producer, or both, can be non-linear in the quantity (or quality) of the object. In contrast, our results thus far have been obtained in a setting with linear utility, and in such environments posted prices are optimal mechanisms. Nonetheless, the same welfare questions can be posed in the general screening environment: what are the feasible pairs of consumer and producer surplus that can be induced through optimal behavior by the monopolist

[8] Even in the discrete case, the supports of prices for consumer surplus maximizing and total surplus minimizing segmentations and pricing rules can only overlap at v^*, the uniform monopoly price.

under some segmentation of the market? While we do not provide a complete answer to this question here, we can report general features of how our results change as we move towards more general screening environments.

In the linear case, the limits of price discrimination are characterized by the surplus triangle, which is defined by the participation constraint of the consumer, the uniform price profit lower bound of the producer, and efficient surplus upper bound. In the non-linear case, there are analogous restrictions on consumer surplus and total surplus, and the monopolist must get at least the *uniform menu profit* that he would obtain with the optimal uniform menu (rather than just a posted price). As we introduce non-linearity, these bounds can no longer be attained exactly, but the central features of the limits survive as follows:

(i) The set of feasible surplus pairs remains "fat" in the sense that many levels of consumer surplus are consistent with profit levels strictly above the uniform menu profit;

(ii) As the non-linear environment approaches linearity, the surplus set of the non-linear environment continuously approaches the surplus triangle of the linear environment.

We will illustrate these points with a simple example that adds a small amount of quadratic concavity to the utility function of the consumer:

$$u_k(q) = v_k q + \epsilon q (1 - q), \quad \epsilon > 0. \tag{7}$$

As before, we maintain zero marginal cost and let $q \in [0, 1]$. The concave model can be interpreted as one of quantity discrimination with a constant marginal cost of production, as in Maskin and Riley (1984). Alternatively, we could have considered a convex cost function to relate the subsequent results to quality discrimination as in Mussa and Rosen (1978). As ϵ goes to zero, the model converges to the linear model in which a uniform price for the entire object, $q = 1$, is always an optimal policy. The concavity in the utility function is independent of the type v, and so is the socially efficient allocation (as long as ϵ is sufficiently small). That is, provided that $v > 0$, the socially efficient allocation is to assign each type a whole unit of the good. In consequence, the efficient boundary of the surplus triangle is independent of the quadratic term and of the size of ϵ.

We explore the consequences of this small amount of concavity with a binary type model: $v_k \in \{1, 2\}$, and the aggregate market is described by the prior probability of the low type:

$$x^* \triangleq \Pr(v_k = 1).$$

This setting will allow us to relate our work to the "concavification" methodology of Aumann and Maschler (1995) and Kamenica and Gentzkow (2011), discussed in the Introduction, which is especially powerful and transparent in the two type case.[9] In the language of Bayesian persuasion (Kamenica and Gentzkow, 2011), the problem is as follows: suppose a sender could commit, before observing his type, to a noisy signal that he will transmit to a receiver conditional on each type. The receiver in turn takes an action with payoff consequences for both sender and receiver, where this action maximizes the receiver's payoff given his posterior beliefs about the type after receiving the signal. This induces a payoff for the sender as a function of the receiver's posterior, and the sender's payoff from the optimal signal structure can be identified by the *concavification* of this payoff function, where the concavification is the smallest concave function that is greater than the sender's payoff. In particular, the maximum payoff of the sender over all signal structures is the concavification evaluated at the prior distribution. In the present problem, the "type" is the consumer's valuation, the "sender" is a social planner who knows the valuation and is maximizing a weighted sum of consumer and producer surplus, and the "receiver" is the monopolist choosing the profit maximizing menu as a function of the posterior belief about the consumer's valuation. In effect, both the signal and the posterior beliefs are a market segment, and Bayesian updating requires that the average distribution across segments equal the aggregate distribution of types. We use this methodology to calculate the surplus frontier by finding the concavification of every possible objective of the social planner.

First, we can solve the monopolist's problem to calculate profit and consumer surplus as a function of x, the proportion of low types. The solution is for the monopolist to exclude the low types if $x < \underline{x}$, to pool the high and low types if $x > \overline{x}$, and to screen by offering the low type an interior allocation $q(x) \in (0,1)$ when $\underline{x} < x < \overline{x}$, where $0 < \underline{x} < \overline{x} < 1$. Thus, profit is:

$$\pi(x) \triangleq \begin{cases} (1-x)\, u_2(1), & \text{if} \quad 0 \leq x \leq \underline{x}; \\ x u_1(q(x)) + (1-x)\left(u_2(1) - u_1(q(x)) + u_2(q(x))\right), & \text{if} \quad \underline{x} \leq x \leq \overline{x}; \\ u_1(1), & \text{if} \quad \overline{x} \leq x \leq 1; \end{cases}$$

[9] In the working paper version, Bergemann, Brooks, and Morris (2013b), we also solve the linear-quadratic utility with a finite number of allocations and a finite number of types, and show how our geometric approach, in particular Lemma 1 has a natural extension to the non-linear environment.

and likewise consumer surplus is:

$$u(x) \triangleq \begin{cases} 0, & \text{if } \ 0 \leq x \leq \underline{x}; \\ (1-x)\left(u_2\left(q\left(x\right)\right) - u_1\left(q\left(x\right)\right)\right), & \text{if } \ \underline{x} \leq x \leq \overline{x}; \\ (1-x)\left(u_2\left(1\right) - u_1\left(1\right)\right), & \text{if } \ \overline{x} \leq x \leq 1. \end{cases}$$

These expressions follow directly from the optimal and incentive compatible menus offered by the seller. The uniform menu profit is $\pi^* = \pi(x^*)$.

In Figure 7, we illustrate the shape of producer and consumer surplus for the quadratic utility function and $q \in [0,1]$ and for $\epsilon = 0.6$, as well as their respective concavified versions. Now, in segments where the concavification is strictly greater than the original function, we find that a convex combination of the critical markets that form the concavification maximizes the given objective. These illustrations immediately indicate some elementary properties of the profit maximizing or consumer surplus maximizing segmentations, which hold true for all concave utility functions $u(v,q)$. The concavified profit function strictly dominates the convex profit function $\pi(x)$ and hence the seller always prefers perfect segmentation, i.e., segments which contain either only low or only high valuations customers. By contrast, it is indicated by the concavified consumer surplus function that the maximal consumer surplus is attained without any segmentation when the share x of low valuation buyers is high, whereas when x is small, market segmentation is required to achieve maximal consumer surplus.

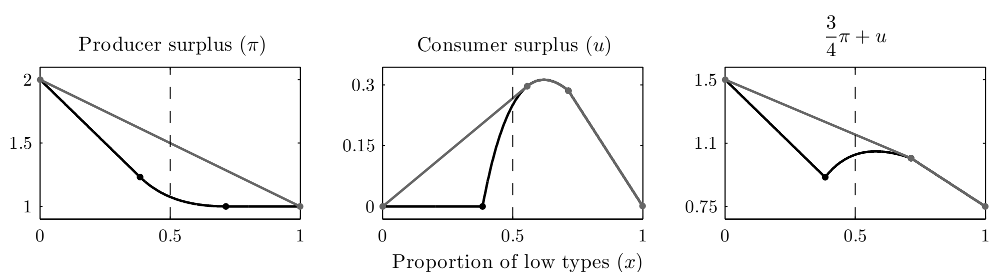

Figure 7: Consumer and producer surplus, and their concavified versions.

The entire boundary of the set of attainable pairs of consumer and producer surplus can be constructed by concavifying the weighted sum of these two expressions, where the weights are allowed to be positive or negative to reach the bounds of the surplus set in all directions. We illustrate the shape of the surplus set in Figure 8 for four different values of ϵ. As ϵ decreases, the concave utility comes closer to the linear utility. At the same time, the surplus set expands and eventually approximates the boundaries that define the surplus triangle, as formally established in Proposition A 1 in the Online Appendix. Importantly, even though for every ϵ there is a unique pair of (u^*, π^*) at the Bayes Nash equilibrium profit level π^*, the overall surplus set is still "fat" in the sense that it has a large interior inside the boundaries identified earlier.

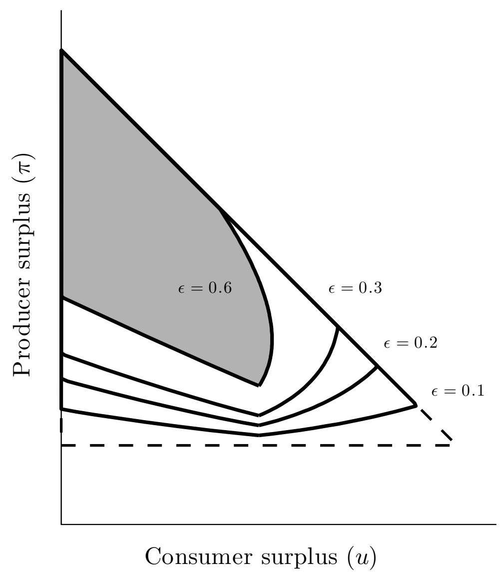

Figure 8: Surplus set with a continuum of allocations.

Ultimately, we find that the dramatic and concise characterization of attainable surplus pairs in our benchmark setting does rely on the linearity of payoffs. Nonetheless, the qualitative features of the main results remain approximately true for small deviations from linearity. The complete characterization of the limits of price discrimination in non-linear settings seems to be an open direction for future research.

6 Conclusion

It was the objective of this paper to study the impact of additional information about consumers' valuations on the distribution of surplus in a canonical setting of monopoly price discrimination. We showed that additional information above and beyond the prior distribution can have a substantial effect on consumer and producer surplus. In general, there are many directions in which welfare could move relative to the benchmark of a unified market. We showed that while additional information can never hurt the seller, it can lead social and consumer surplus to both increase, both decrease, or respectively increase and decrease. Most notably, we establish the exact limits of these predictions without any restrictions on the aggregate market, and in particular, the sharp boundaries do not rely on any regularity or concavity assumption on the distribution of values or the profit function.

Exactly which form of market segmentation arises in practice is no doubt influenced by many factors, which may include technological and legal limitations on how information can be collected and used. In an age in which individuals are increasingly concerned about the preservation of privacy, it is important to understand the welfare consequences that may result from the collection of data on consumers' preferences. Policy discussion often assumes that this will favor producers and hurt consumers. This may be a reasonable assumption if the data ends up in the hands of producers. But this need not be the case. Our findings indicate that the relationship between efficiency and information can only be understood in the context of how data will be used, and this crucially depends on the preferences of those who collect the information. Thus, a natural and important direction for future research is to better understand which forms of price discrimination will endogenously arise, and for whose benefit.

References

Aguirre, I., S. Cowan, and J. Vickers (2010): "Monopoly Price Discrimination and Demand Curvature," *American Economic Review*, 100, 1601–1615.

Aumann, R., and M. Maschler (1995): *Repeated Games with Incomplete Information*. MIT.

Bergemann, D., B. Brooks, and S. Morris (2013a): "Extremal Information Structures in First Price Auctions," Discussion Paper 1926, Cowles Foundation for Research in Economics, Yale University.

——— (2013b): "The Limits of Price Discrimination," Discussion Paper 1896RR, Cowles Foundation for Research in Economics, Yale University.

Bergemann, D., and S. Morris (2013a): "The Comparison of Information Structures in Games: Bayes Correlated Equilibrium and Individual Sufficiency," Discussion Paper 1909, Cowles Foundation for Research in Economics, Yale University.

——— (2013b): "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81, 1251–1308.

Blackwell, D. (1951): "Comparison of Experiments," in *Proc. Second Berkeley Symp. Math. Statist. Probab.*, pp. 93–102. University of California Press, Berkeley.

——— (1953): "Equivalent Comparison of Experiments," *Annals of Mathematics and Statistics*, 24, 265–272.

Cowan, S. (2013): "Welfare-Increasing Third-Degree Price Discrimination," Discussion paper, Oxford University.

Kamenica, E., and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

Mas-Colell, A., M. Whinston, and J. Green (1995): *Microeconomic Theory*. Oxford University Press.

Maskin, E., and J. Riley (1984): "Monopoly with Incomplete Information," *Rand Journal of Economics*, 15, 171–196.

Mussa, M., and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

Nahata, B., K. Ostaszewski, and P. Sahoo (1990): "Direction of Price Changes in Third-Degree Price Discrimination," *American Economic Review*, 80, 1254–1258.

Pigou, A. (1920): *The Economics of Welfare*. Macmillan, London.

Riley, J., and R. Zeckhauser (1983): "Optimal Selling Strategies: When to Haggle, When to Hold Firm," *Quarterly Journal of Economics*, 98, 267–290.

Robinson, J. (1933): *The Economics of Imperfect Competition*. Macmillan, London.

Schmalensee, R. (1981): "Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination," *American Economic Review*, 71, 242–247.

Simon, B. (2011): *Convexity: An Analytic Viewpoint*. Cambridge University Press.

Varian, H. (1985): "Price Discrimination and Social Welfare," *American Economic Review*, 75, 870–875.

A Proofs

Proof of Lemma 2. We begin with the following four observations.

Fact 1: LC implies that optimality constraints are satisfied for all $i \geq l + 1$.

Fact 2: UC implies that the non-negativity constraint is slack for some $i \geq l$.

Fact 3: If LC binds, then the non-negativity constraint is slack for some $i \geq l + 1$.

Fact 4: If UC binds and $x_i > 0$ for some $i > l$, then v_l is not optimal.

As with X_k, any element $x \in X_{k,l}$ must satisfy the $2(K-1)$ non-negativity and optimality constraints. In addition, x must satisfy the lower and upper minimum output constraints. An extreme point of $X_{k,l}$ is characterized by a subset of at least $K-1$ of these constraints which are binding (Simon, 2011, Proposition 15.2). By Fact 1, we can drop all of the optimality constraints for $i \geq l+1$, leaving just UC, LC, non-negativity, and optimality constraints for $i \leq l$. Let $m = |\{l, \ldots, K\} \setminus \{k\}|$. As before, non-negativity of x_i and optimality of v_i are mutually exclusive, so we can obtain at most $K-1-m$ binding optimality or non-negativity constraints for $i < l$, leaving m constraints to define the extreme point. We will establish by cases that the choice of the remaining active constraints must be equivalent to choosing at most one optimality constraint from $\{l+1, \ldots, K\}$ and at least one optimality constraint from $\{l, \ldots, K\}$, with the remaining binding constraints being non-negativity. Note that if $l = K$, then LC is always satisfied and can be dropped, and UC implies that v_K is an optimal price.

First, suppose that non-negativity is binding for x_l, i.e., $x_l = 0$. Note that this case can only arise if $l < K$. Then UC and LC are redundant and equivalent to the single linear restriction:

$$v^* \sum_{j=i^*}^{K} x_j = \sum_{j=l+1}^{K} v_j x_j.$$

By Fact 2, there must be exactly $m-1$ binding non-negativity constraints for $i \geq l$. This implies that optimality binds for the single $i \geq l+1$ for which the non-negativity constraint on x_i need not bind.

Now suppose that $x_l > 0$. Then at most one of UC and LC can bind (the gap between them is strict). Suppose that it is LC. By Fact 3, some non-negativity constraint is slack for $i \geq l+1$. If more than one of these is slack, or if optimality is not binding for v_l, we could not reach the requisite number of binding constraints. Thus, exactly one non-negativity constraint is slack for $i \geq l+1$, and optimality binds for v_l. Together with LC, this gives us the required m. As a result, optimality must be binding for the single $i \geq l+1$ for which non-negativity does not bind.

Alternatively, suppose $x_l > 0$ and UC binds. (Note that this case cannot arise when $l = k$, since then UC is implied by optimality of v_k, and can be dropped from the problem.) By Fact 4, non-negativity must bind for all $i \geq l+1$. Otherwise, neither optimality of v_l, non-negativity of x_l, nor non-negativity of x_i for $i \geq l+1$ binds, and we would only have UC plus at most $m-2$ non-negativity constraints for $i \geq l+1$, which is one short of the m required. We conclude that $x_i = 0$ for all $i \geq l+1$, so that UC implies that optimality binds for v_l.

Finally, it must be that either LC or UC binds. Otherwise, then clearly $x_l > 0$ and the only way to get to $K-1$ constraints is all non-negativity constraints are binding for $i \geq l+1$, and optimality binding for v_l, which contradicts that LC does not bind. □

Proof of Lemma 3. As X is metrizable in the Prokhorov metric, $\operatorname{cl} \widehat{X}_v$ is the set of limits of convergent sequences in $\widehat{X}_v$. Take a weakly convergent sequence $x_k \in \widehat{X}_v$, and suppose that it converges to some $x \notin \overline{X}_v$. Then we can find some $v' \in \operatorname{supp} x$ such that $v'\, x\left([v', \overline{v}]\right) \neq \min \operatorname{supp} x$.

Claim: For any $v' \in \operatorname{supp} x$ and for any $\epsilon > 0$, there exists a K such that for all $k \geq K$, we can find $v_k \in \operatorname{supp} x_k$ such that $|v_k\, x\left([v_k, \overline{v}]\right) - v'\, x\left([v', \overline{v}]\right)| < \epsilon$. The measure x can have at most countably many mass points, so we can find a $\delta > 0$ so that $x\left(\{v' - \delta, v' + \delta\}\right) = 0$ (i.e., both $[v' - \delta, v' + \delta]$ and $[v' - \delta, \overline{v}]$ are continuity sets of x), $x\left([v' - \delta, v')\right) < \epsilon/(3\overline{v})$, and $\delta \leq \epsilon/3$. By weak convergence, we must have that $x_k\left([v' - \delta, \overline{v}]\right) \to x\left([v' - \delta, \overline{v}]\right)$ and $x_k\left([v' - \delta, v' + \delta]\right) \to x\left([v' - \delta, v' + \delta]\right)$. Since $v' \in \operatorname{supp} x$, it must be that $x\left([v' - \delta, v' + \delta]\right) > 0$, so we can pick a K large enough so that for all $k \geq K$, $x_k\left([v' - \delta, v' + \delta]\right) > 0$ and $|x_k\left([v' - \delta, \overline{v}]\right) - x\left([v' - \delta, \overline{v}]\right)| < \epsilon/(3\overline{v})$. Let $v_k = \min \operatorname{supp} x_k \cap [v' - \delta, v' + \delta]$, which is non-empty because $x_k\left([v' - \delta, v' + \delta]\right) > 0$. Hence,

$$\begin{aligned}
|v_k\, x_k\left([v_k, \overline{v}]\right) - v\, x\left([v', \overline{v}]\right)| &= |v_k\, x_k\left([v' - \delta, \overline{v}]\right) - v\, x\left([v', \overline{v}]\right)| \\
&\leq \left| v_k\, x_k\left([v' - \delta, \overline{v}]\right) - v'\, x\left([v' - \delta, \overline{v}]\right) + v' \frac{\epsilon}{3\overline{v}} \right| \\
&\leq \left| v_k \frac{\epsilon}{3\overline{v}} \right| + |v_k - v'| + \left| v' \frac{\epsilon}{3\overline{v}} \right| \\
&\leq \epsilon,
\end{aligned}$$

which proves the claim.

Thus, we can find K large enough so that for $k \geq K$, there exist v_k and $v'_k \in \operatorname{supp} x_k$ such that $|v_k\, x_k\left([v_k, \overline{v}]\right) - \min \operatorname{supp} x| < \epsilon$ and $|v'_k\, x_k\left([v'_k, \overline{v}]\right) - v'\, x\left([v', \overline{v}]\right)| < \epsilon$ where $\epsilon = \left(|v'\, x\left([v', \overline{v}]\right) - \min \operatorname{supp} x|\right) /$ But this means that $|v_k\, x_k\left([v_k, \overline{v}]\right) - v'_k\, x_k\left([v'_k, \overline{v}]\right)| > 0$, which contradicts the assumption that $x_k \in \widehat{X}_v$. □

Proof of Proposition 1B. The monopolist is indifferent to a pricing rule that puts probability one on v^* when $x \in \overline{X}_{v^*}$. Under such a rule, producer surplus is π^*. By definition of an extremal segmentation, the maximum and minimum pricing rules are optimal, and the former results in a consumer surplus of zero and the latter results in all consumers purchasing the good, so that total surplus is w^*. $\square$

Proof of Theorem 1B. The argument for necessity is as in the finite case. For sufficiency, we first argue that there exists a uniform profit preserving extremal segmentation. Since simple measures are dense in X_{v^*}, we can find a sequence x_k of markets in X_{v^*} that converge to x^* in the weak topology. By Lemma 1, there exist extremal segmentations of x_k for every k, which we can identify with elements σ_k of $\Delta\left(\operatorname{cl}\widehat{X}_{v^*}\right)$. Since $\operatorname{cl}\widehat{X}_{v^*}$ is a closed subset of the compact set X, it is also a compact. By the Banach-Alaoglu Theorem, $\Delta\left(\operatorname{cl}\widehat{X}_{v^*}\right)$ is compact, so σ_k has a convergent subsequence that converges to some $\sigma \in \Delta\left(\operatorname{cl}\widehat{X}_{v^*}\right)$, and therefore $\sigma \in \Delta\left(\overline{X}_v\right)$ as well, i.e., σ is a uniform profit preserving extremal segmentation.

Claim: σ is a segmentation of x^*. For any continuous and bounded function f on $[0,\overline{v}]$, we have

$$\begin{aligned}
\int_{v\in[0,\overline{v}]} f(v)\,x(dv) &= \lim_{k\to\infty}\int_{v\in[0,\overline{v}]} f(v)\,x_k(dv)\\
&= \lim_{k\to\infty}\int_{v\in[0,\overline{v}]} f(v)\int_{x\in X} x(dv)\,\sigma_k(dx)\\
&= \lim_{k\to\infty}\int_{x\in X}\int_{v\in[0,\overline{v}]} f(v)\,x(dv)\,\sigma_k(dx)\\
&= \int_{x\in X}\int_{v\in[0,\overline{v}]} f(v)\,x(dv)\,\sigma(dx),
\end{aligned}$$

where the first line follows from weak convergence, the second line is the definition of an extremal segmentation, the third line is Fubini's Theorem, and the last line is again weak convergence, using the fact that $\int_{v\in[0,\overline{v}]} f(v)\,x(dv)$ is a continuous function of x. Hence, the measure $\int_{x\in X} x\cdot\sigma(dx)$ is a version of x^*.

To conclude, there exists a uniform profit preserving segmentation σ, under which minimum and maximum pricing rules are optimal and induce the points $(w^*-\pi^*,\pi^*)$ and $(0,\pi^*)$. As before, the segmentation σ' defined by $\sigma'(Y) = x^*\left(\left\{v \middle| x^{\{v\}}\in Y\right\}\right)$ and $x^{\{v\}}$ is the Dirac measure on $\{v\}$, together with any optimal pricing rule, induces the welfare outcome $(0,w^*)$. Weighted averages $\sigma'' = \alpha\sigma' + (1-\alpha)\sigma$ and $\phi = \beta\underline{\phi} + (1-\beta)\overline{\phi}$ achieve every other surplus pair in the triangle. $\square$

B Non-linear Analysis

This Appendix provides a formal analysis of the example of Section 5. We observe that optimal menus have $q_2 = 1$, and thus we define the quantity of low type for simplicity as $q_1 \triangleq q$, and as individual rationality binds for the low type, it must be that $p_1 = q + \epsilon q (1 - q)$. The benefit of the high type from pretending to be the low type is thus:

$$2q + \epsilon q (1 - q) - p_1 = q.$$

Since the high type has a binding incentive constraint, his payment is simply $p_2 = 2 - q$. When there is a proportion $x_1 \triangleq x$ of low types, profits when allocating q to the low type are:

$$(q + \epsilon q (1 - q)) x + (2 - q)(1 - x).$$

Differentiating with respect to q, the first order condition is:

$$0 = (1 + \epsilon (1 - 2q)) x - (1 - x) \Leftrightarrow q(x) = \frac{1}{2} - \frac{1 - 2x}{2\epsilon x}.$$

If this number is in $[0, 1]$, then the optimal menu sells this quantity to the low type. If it is negative, for which the condition is:

$$\frac{1}{2} \leq \frac{1 - 2x}{2\epsilon x} \Leftrightarrow x \leq \frac{1}{2 + \epsilon},$$

then since profits are concave in q, it is optimal to exclude the low type with an allocation of $q = 0$. Note that when $x < \frac{1}{2}$, excluding the low type also yields a strictly higher payoff than pooling.

Similarly, if the solution to the first-order condition is greater than 1, for which the condition is:

$$-\frac{1}{2} \geq \frac{1 - 2x}{2\epsilon x} \Leftrightarrow x \geq \frac{1}{2 - \epsilon},$$

then it is optimal to pool the high type and the low type at the efficient output. Note that when $x > \frac{1}{2}$, pooling yields strictly higher profit than exclusion. Hence, output is:

$$q(x) = \begin{cases} 0, & \text{if} \quad 0 \leq x < \frac{1}{2+\epsilon}; \\ \frac{1}{2} - \frac{1-2x}{2\epsilon x}, & \text{if} \quad \frac{1}{2+\epsilon} \leq x < \frac{1}{2-\epsilon}; \\ 1, & \text{if} \quad \frac{1}{2-\epsilon} \leq x \leq 1. \end{cases}$$

Producer surplus is given by:

$$\pi(x) = q(x)(1 + \epsilon (1 - q(x))) x + (2 - q(x))(1 - x);$$

and consumer surplus is given by:

$$u(x) = (1 - x)\, q(x),$$

which are illustrated in Figure 7.

We will solve for the surplus set when the aggregate market is $x^* = \frac{1}{2}$. We can write:

$$w_\lambda(x) \triangleq \lambda\pi(x) + u(x)$$

for $\lambda \in \mathbb{R}$. The support function of the surplus set at a given direction $(1, \lambda)$ is given by the concavification of $w_\lambda(x)$ at x^*. Similarly, the support function at directions $(-1, -\lambda)$ is given by the concavification of $-w_\lambda(x)$ at x^*.

Eastern Frontier The concavification of w_λ at x^* falls into four "regimes", defined relative to three cutoff values of λ which are $\underline{\lambda} < \widehat{\lambda} < \overline{\lambda}$. We depict examples for each of these regimes in Figure 9. For each of four values of λ, we plot scaled producer surplus, consumer surplus, the sum, as well as their respective concavifications. The plots are scaled to show the relative magnitudes of $\lambda\pi$ and u, with λ decreasing as we progress downwards through the figure. For λ extremely large, i.e., $\lambda \in \left(\overline{\lambda}, \infty\right)$, we are close to maximizing π. This is accomplished by perfect price discrimination, in which the market is segmenting into $x = 0$ and $x = 1$, as depicted in the top row of Figure 9. For this range of λ, the extreme point of the welfare set is the efficient point where $u = 0$, which the northernmost point in Figure 8.

As λ decreases, consumer surplus becomes more important, and the concavification of w_λ comes closer to the line segment between $x = 1/(2 - \epsilon)$ and $x = 1$, where both u and π are linear. At a critical value $\overline{\lambda}$, the two coincide, in particular when:

$$\frac{1}{2 - \epsilon} w_\lambda(1) + \left(1 - \frac{1}{2 - \epsilon}\right) w_\lambda(0) = w_\lambda\left(\frac{1}{2 - \epsilon}\right).$$

The solution to this equation is given by $\overline{\lambda} = 1$. At this point, the concavification at $x = 1/2$ becomes the line segment connecting $(0, w_\lambda(0))$ and $(1/(2 - \epsilon), w_\lambda(1/(2 - \epsilon)))$. In other words, for the next range of directions, it is optimal to segment between markets with $x = 0$ and $x = 1/(2 - \epsilon)$. This case is depicted in the second row of Figure 9, where $\lambda = 3/4$. Note that as λ decreases, this corresponds to the direction we are maximizing in rotating clockwise from due north. At some point, the optimum switches from the northernmost point to the easternmost point on the efficient frontier in Figure 8.

As λ decreases further towards zero, $w_\lambda(0) = 2\lambda$ falls relative to w_λ in $[1/(2 + \epsilon), 1/(2 - \epsilon)]$. Eventually, the tangent between $(0, w_\lambda(0))$ and the graph of w_λ moves from being at $1/(2 - \epsilon)$ to a

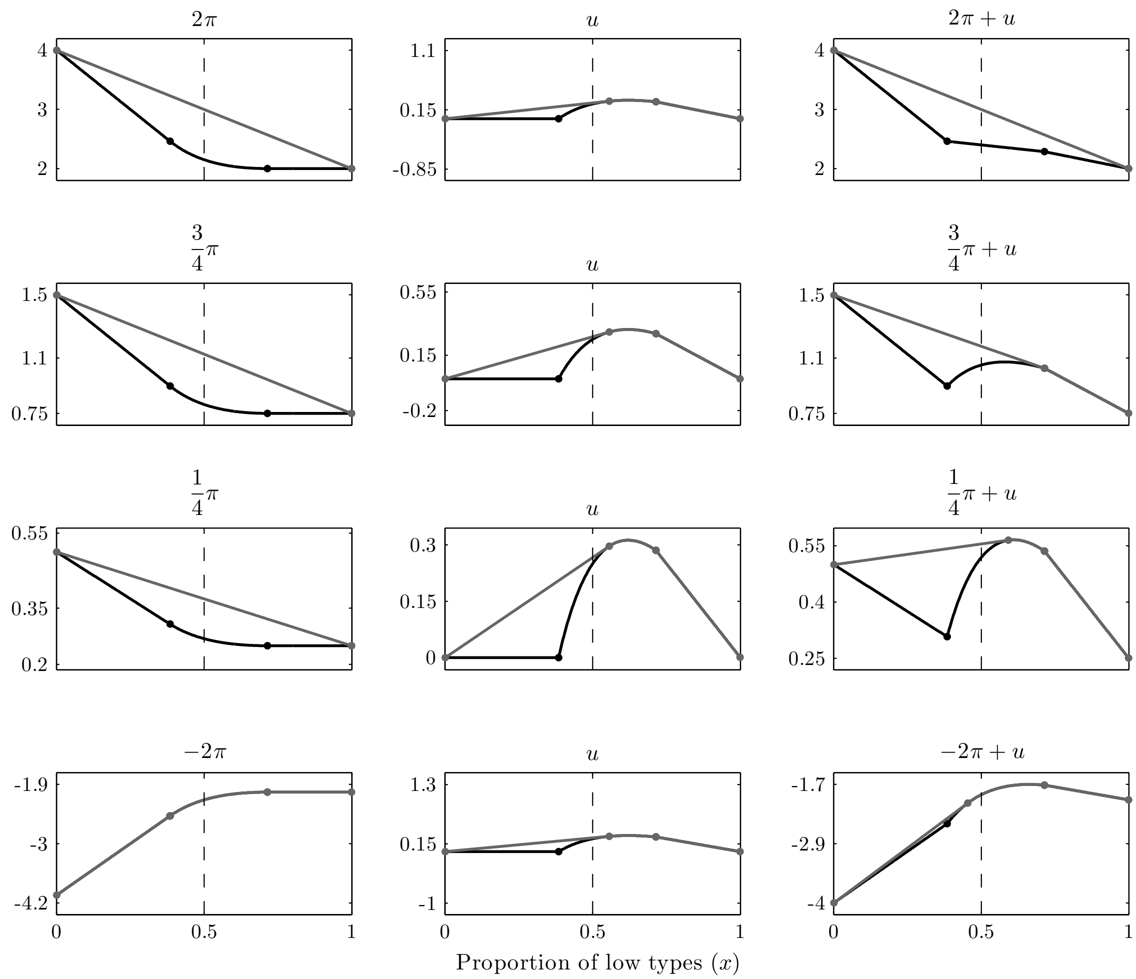

Figure 9: Examples for the concavification along the eastern frontier.

point in $[1/2, 1/(2-\epsilon)]$. The tangent just moves to the left of $1/(2-\epsilon)$ when:

$$\frac{1}{2-\epsilon} \lim_{x \uparrow \frac{1}{2-\epsilon}} w_\lambda'(x) = \left(w_\lambda \left(\frac{1}{2-\epsilon} \right) - w_\lambda(0) \right).$$

The solution is:

$$\widehat{\lambda} = \frac{3}{2} - \frac{1}{2\epsilon}.$$

In this regime, the concavification at $x = 1/2$ is given by the line connecting $(0, w_\lambda(0))$ and $(x(\lambda), w_\lambda(x(\lambda)))$, where:

$$x\, w_\lambda'(x) = w_\lambda(x) - w_\lambda(0) \Leftrightarrow x(\lambda) = \frac{2-\lambda}{3 + \epsilon(1-\lambda) - 2\lambda}.$$

This is the case in the third row of Figure 9, where $\lambda = 1/4$. For λ in this range, we trace out the curved portion of the eastern-southeastern frontier of Figure 8. The final cutoff $\underline{\lambda}$ is the solution to:

$$\frac{1}{2}w_\lambda'\left(\frac{1}{2}\right) = w_\lambda\left(\frac{1}{2}\right) - w_\lambda(0) \Leftrightarrow \underline{\lambda} = 1 - \frac{1}{\epsilon}.$$

It is at this point that the tangent point from $(0, w_\lambda(0))$ to the graph of w_λ moves to the left of $1/2$, so that the concavification of w_λ at $x = 1/2$ is just $w_\lambda(1/2)$, which is true for all $\lambda \in (-\infty, \underline{\lambda})$. In this range, the weight on minimizing producer surplus is so large that the solution is no segmentation, as in the bottom row of Figure 9. For these values, the direction $(1, \lambda)$ points southerly enough that the optimum is no information, i.e., the southern corner of Figure 8.

Western Frontier For the western frontier, we find the concavification of $-w_\lambda(x)$ at $x = 1/2$ for $\lambda \in \mathbb{R}$. As before, there are four regimes, and we depict examples in Figure 10. For λ sufficiently large, again we are close to minimizing π, and no segmentation is optimal, as in the top row. Note that large λ corresponds to maximizing a direction $(-1, -\lambda)$ close to due south, so again we are at the southern corner of Figure 8.

As λ decreases, the weight on u becomes relatively large compared to the weight on π, and eventually the concavification at $x = 1/2$ switches to the tangent line between $(1, w_\lambda(1)) = (1, \lambda)$ and the graph of $w_\lambda(x)$. Let $x(\lambda)$ again denote the point of tangency, which solves:

$$(1-x)\,w_\lambda'(x) = 2\lambda - w_\lambda(x) \implies x(\lambda) = 1/\left(1 - \frac{\sqrt{(1-\epsilon(6-\epsilon))\,\lambda(\lambda-2)}}{\lambda-2}\right).$$

The critical $\underline{\lambda}$ occurs when $x(\lambda) = 1/2$, which is:

$$\underline{\lambda} = \frac{2}{\epsilon(6-\epsilon)}.$$

As λ increases above $\underline{\lambda}$, $x(\lambda)$ decreases from $1/2$ until it eventually hits $1/(2+\epsilon)$. The critical λ is:

$$\widehat{\lambda} = \frac{(1+\epsilon)^2}{4\epsilon}.$$

For $\lambda \in \left[\underline{\lambda}, \widehat{\lambda}\right]$, the solution is to segment using markets $x(\lambda)$ and $x = 1$, as in the second row of Figure 10. Thus, in Figure 8, there is in fact a subtle curve to the southwestern frontier as $x(\lambda)$ moves smoothly from $1/2$ to $1/(2+\epsilon)$. At $\widehat{\lambda}$, the regime changes to segmenting between $x = 1/(2+\epsilon)$ and $x = 1$, as in the third row of Figure 10. This generates the southernmost point along the western frontier where $u = 0$. This continues until we hit $\overline{\lambda}$ at which:

$$w_\lambda\left(\frac{1}{2+\epsilon}\right) = \frac{1}{2+\epsilon}w_\lambda(1) + \left(1 - \frac{1}{2+\epsilon}\right)w_\lambda(0),$$

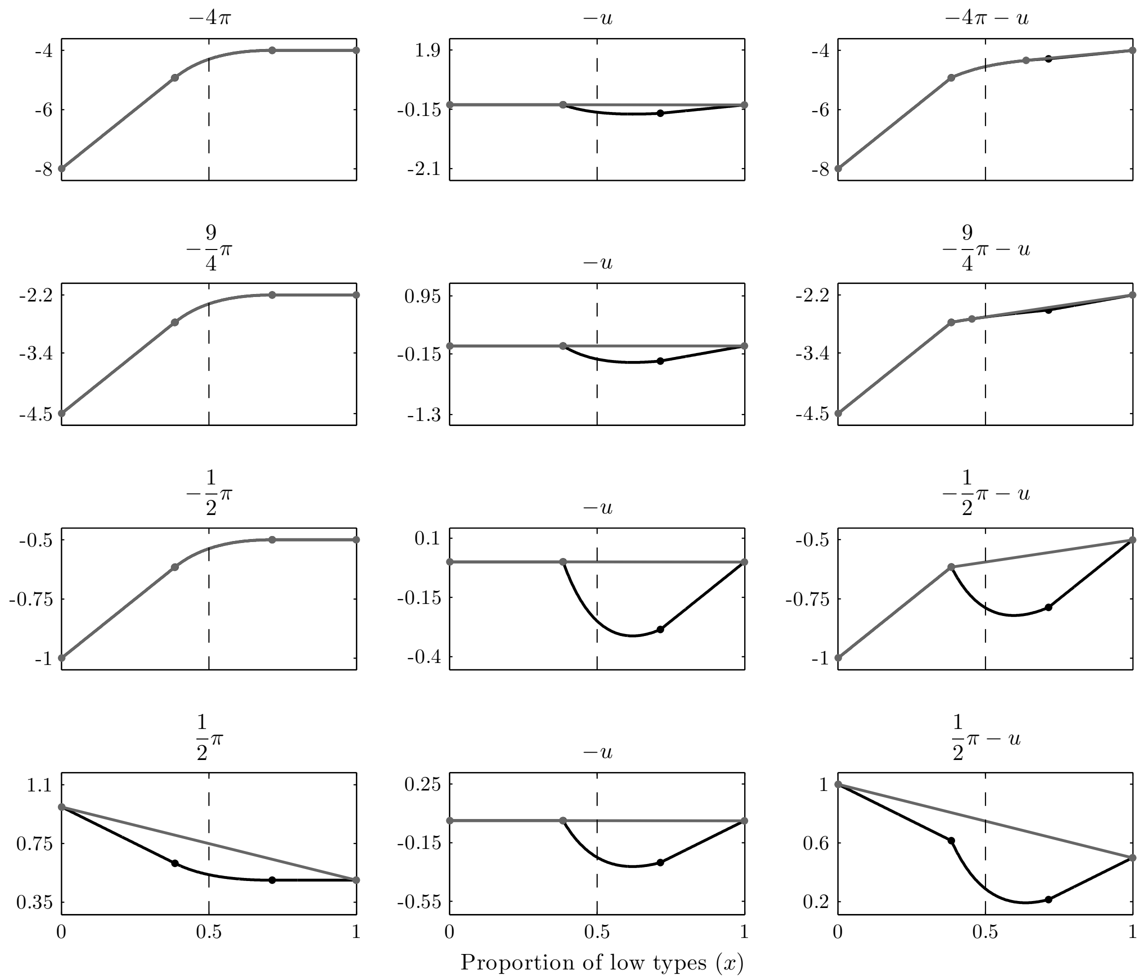

Figure 10: Examples for the concavification along the western frontier.

which occurs precisely at $\overline{\lambda} = 0$, when we are minimizing consumer surplus. Finally, for $\lambda \in \left(\overline{\lambda}, \infty\right)$, when we have negative weight on consumer surplus and a non-negative weight on producer surplus, the optimum is again perfect price discrimination, and we are back to the northernmost corner of Figure 8.

Limit as $\epsilon \to 0$ This characterization of the frontier can be used to study the limit of the surplus set as $\epsilon \to 0$. In particular, we know that for ϵ sufficiently small, there are directions for which segmenting $x^* = 1/2$ into $\{0, 1\}$, $\{0, 1/(2-\epsilon)\}$, and $\{1/(2+\epsilon), 1\}$ are respectively optimal.

For every ϵ, segmentation into $\{0, 1\}$ induces consumer surplus of zero and producer surplus equal to $3/2$, so this is also the limit as $\epsilon \to 0$.

In the limit as $\epsilon \to 0$, the segmentation of $1/2$ into $\{0, 1/(2-\epsilon)\}$ puts probability approaching 1 on the segment with $x = 1/(2-\epsilon)$. In this segment, the monopolist is indifferent between screening and selling the efficient quantity to both types at a price of 1. We break indifference in favor of the latter so that the monopolist receives a profit of 1 and the allocation is efficient. Asymptotically, this segment dominates so that total surplus is efficient and profit is $\pi^* = 1$.

In the limit as $\epsilon \to 0$, the segmentation of $1/2$ into $\{1/(2+\epsilon), 1\}$ puts probability approaching 1 on the segment with $x = 1/(2+\epsilon)$. In this segment, the monopolist only sells to the high type and garners a profit of $2(1-\epsilon)/(2+\epsilon)$, which is converging to 1 as $\epsilon \to 0$. In the limit, consumer surplus is 0 and profit is $\pi^* = 1$.

Finally, we briefly comment on what happens in the limit for $x^* \neq 1/2$. If $x^* > 1/2$, then for ϵ sufficiently small it is optimal to pool in the aggregate market, so that no segmentation exactly achieves point C in Figure 1. Point D can be attained by segmenting into $\{1/(2+\epsilon), 1\}$, and having the monopolist "pool" in the market with $x = 1$ and exclude in the market with $1/(2+\epsilon)$ and sell only to the high types. The weight on the market with $x = 1/(2+\epsilon)$ is $(1-x^*)(2+\epsilon)/(1-\epsilon)$, so that the monopolist's profit is:

$$\frac{2+\epsilon}{1-\epsilon}(1-x^*)2\frac{1}{2+\epsilon} + 1 - \frac{2+\epsilon}{1-\epsilon}(1-x^*),$$

which approaches 1 as $\epsilon \to 1$. Similarly, if $x^* < 1/2$, then it is optimal to exclude in the aggregate market for ϵ sufficiently small, so no segmentation achieves D. And by a similar calculation as before, segmentation into $\{0, 1/(2-\epsilon)\}$ achieves point C in the limit.

Thus, the three extreme points of the surplus triangle are attained in the limit, and by convexity the entire triangle must be attained as well. We can then summarize these results as follows (and as reflected in Figure 8).

Proposition A5 (Surplus Set Close to Linearity).
In the linear-quadratic model (7), the set of attainable profit and consumer surplus pairs converges to the surplus triangle of the uniform price monopoly as $\epsilon \to 0$.